Rev Manag Sci (2016) 10:511–552
DOI 10.1007/s11846-015-0166-8

ORIGINAL PAPER

Overconfidence and risk seeking in credit markets: an experimental game

David Peón · Manel Antelo · Anxo Calvo

Received: 3 April 2014 / Accepted: 5 February 2015 / Published online: 4 March 2015

Abstract Behavioral biases may influence bank decisions when granting credit to their customers. This paper explores this possibility in an experimental setting, contributing to the literature in two ways. First, we designed a business simulation game that replicates the basic decision-making processes of a bank granting credit to clients under conditions of risk and uncertainty. Second, we implemented a series of short tests to measure participants' overconfidence and risk profile according to prospect theory and then conduct an experimental implementation of the simulation game. We find that higher levels of overprecision and risk seeking for gains (mostly attributable to distortion of probabilities) foster lower prices and higher volumes of credit, and reduce quality. The most consistent result is that distortion of probabilities affects the ability to discriminate between the quality of borrowers according to objective information, fostering strategies of lower loan prices to lower quality clients. The external validity of the results is also discussed.

Electronic supplementary material The online version of this article (doi:10.1007/s11846-015-0166-8) contains supplementary material, which is available to authorized users.

D. Peón
Grupo BBVA, A Coruña, Spain

D. Peón (✉) · A. Calvo
Department of Financial Economics and Accountancy, University of A Coruna,
Campus Elviña s/n, 15071 A Coruña, Spain
e-mail: david.peon@udc.es

A. Calvo
e-mail: anxo.calvo@udc.es

M. Antelo
Department of Economics, University of Santiago de Compostela, Campus Norte,
15782 Santiago de Compostela, Spain
e-mail: manel.antelo@usc.es

Keywords Experimental economics · Business simulation games · Banking · Credit markets · Overconfidence · Prospect theory

JEL Classification C91 · D03 · D81 · G21

1 Introduction

The good performance of western economies before the Great Recession, largely fueled by credit, came to a dramatic end with the financial crisis of 2008. Credit crises have been traditionally associated with both credit-supply and credit-demand effects (Presbitero 2012). On the supply side, researchers have analyzed, among others, the role incentives (Fahlenbrach and Stulz 2011), securitization (Keys et al. 2010) and moral hazard risk-taking by banks (Acharya and Naqvi 2012) might have played in the current crisis. Behavioral economics may offer a complementary interpretation; in particular, the effect of behavioral biases by participants in the banking industry might be helpful to explain how credit booms are fueled by the banking sector.

In this research we focus on two of these behavioral biases: overconfidence and prospect theory. Overconfidence may manifest itself in several ways. Moore and Healy (2008) offer a classic taxonomy: we may be overconfident in estimating our own performance (known as *overestimation*), in estimating our own performance relative to others (*overplacement* or better-than-average effect), or in applying excessive precision in estimating future uncertainty (known as *overprecision*). Prospect theory, today the most widely accepted descriptive decision theory, followed after Kahneman and Tversky's (1979) experimental evidence that people, when evaluating prospects in order to make a decision, tend to treat gains differently to losses and to overweight outcomes with small probabilities.

The effects of overconfidence and prospect theory are well-known in the literature. Overconfidence has been claimed to explain anomalies like excess volatility, under- and overreaction (Daniel et al. 1998), excessive trading (Odean 1998, 1999), asset bubbles (Scheinkman and Xiong 2003), and the forward premium puzzle (Burnside et al. 2011). Risk seeking would explain anomalies like the house money effect (Thaler and Johnson 1990) and the status quo bias (Tversky and Kahneman 1991). Probability weighting explains the favorite-longshot bias and portfolio under-diversification (Hens and Bachmann 2008; Barberis and Huang 2008), while loss aversion would explain several anomalies in decision making, like the endowment effect (Thaler 1980), the disposition effect (Shefrin and Statman 1985), the status quo bias (Samuelson and Zeckhauser 1988), the equity premium puzzle (Benartzi and Thaler 1995), and how the number of transactions in the market would be reduced (Knetsch 1989).

The effects on managerial performance of overconfidence and risk profile according to prospect theory are also well known. Executives appear to be particularly prone to displaying overconfidence (Moore 1977), which helps explain the high rates of business failure (Camerer and Lovallo 1999), high rates of corporate merger and acquisition (Roll 1986; Malmendier and Tate 2005a, b), lower dividend payouts (Deshmukh et al. 2010), and higher cash holdings (Huang et al. 2012), among other anomalies in corporate finance. Besides, different manifestations of prospect theory, such as probability weighting, aversion to a sure loss and lower loss aversion, would help explain the IPO underpricing puzzle (Barberis and Huang 2008), risky capital budgeting decisions (Shefrin 2008; Shefrin and Cervellati 2011) and why managers would take more risks (Rabin 2000), respectively.

Several experiments confirmed that judgmental overconfidence has an impact on financial decision making. Thus, Biais et al. (2005) and Grinblatt and Keloharju (2009) confirm the negative effects of overconfidence on the profitability of trading suggested by Barber and Odean (2001, 2002), while Glaser and Weber (2007) emphasize the need to consider different types of confidence, as different measures of miscalibration and the better-than-average effect often yield conflicting results. In addition, a recent and growing research field is the analysis of credit cycles, in which several behavioral models have been proposed that are based on overconfidence effects (Keen 2011; Rötheli 2012; Peón et al. 2015). Finally, some research has also been performed regarding financial professionals behaving according to prospect theory (Abdellaoui et al. 2013) and how personal characteristics of executive teams might affect corporate governance in banking (Berger et al. 2012).

The main goal of this paper is to obtain experimental evidence of these behaviorally driven effects on retail credit markets. In particular, we trace whether overconfidence and different prospect theory profiles by participants in the banking industry could feed risk-seeking behavior that explains, to some extent, excessive lending by retail banks. For that purpose, we designed an original business simulation game that replicates the basics of a bank granting credit to costumers, and organized a series of five experimental sessions with students in the University of A Coruna (Spain) in 2013 to test the effects of different levels of overconfidence and risk profile—according to prospect theory—on the credit policies implemented in the experiment. In total, 126 volunteers, all of them under- and postgraduate students, participated in the experiment. Before participation, we determined the psychological profile (based on overconfidence and prospect theory) of each participant, through a series of short tests implemented following Peón et al. (2014). The students then participated in the strategy game designed to replicate, in an experimental setting, how banks grant credit to costumers, in order to obtain information about how much and at what price different subjects would grant credit under conditions of uncertainty and risk about the economic environment.

Our main contributions are two. First, we designed a business simulation game that replicates the basics of the decision-making process of a bank granting credit to costumers under conditions of uncertainty and risk. We review the literature on simulation games, banking efficiency and credit markets to justify a game design that meets the required characteristics of simulations (Gredler 2004). Thus, the design imitates a complex situation where objective information provided to participants about economic perspectives and customers' expected solvency is presented in the form of confidence intervals, and is updated period by period to introduce a feedback system that allows participants to learn how economic perspectives evolved. With the cost structure as an input, participants are required to set their strategies every period in terms of price and volume of loans granted to each niche of clients. Two variables are also introduced to control for loan quality.

The second contribution is the experimental implementation of the game together with a series of psychological tests to measure participants' level of overconfidence and risk profile according to prospect theory. This allowed us to test the effects of their behavioral biases on the credit policies they implement in the game. Monetary incentives were introduced to improve the external validity of the results. Several hypotheses about the effects of risk seeking, loss aversion and overconfidence were tested. Experimental results provide extensive evidence that more aggressive behavioral profiles—in terms of higher overconfidence and risk seeking—are correlated with more aggressive credit strategies. The most conclusive results are greater overprecision and distortion of probabilities in the positive domain fostering riskier credit strategies, particularly in terms of providing credit to low-quality costumers at a lower price. All the statistical analyses (correlations, regressions, factorial and clustering analyses) provide consistent results in that sense.

The remainder of the article is organized as follows. First, in Sect. 2 we discuss how the experiment was designed, the basics of the game and the variables to be measured. Section 3 describes the hypotheses to be tested. Section 4 analyzes the main results obtained. Section 5 concludes. Additional statistical information and technical specifications of the statistical analyses are provided in the Electronic Supplementary Material.

2 Experiment design

We organized a series of five experimental sessions that took place in the Faculty of Business and Economics of the University of A Coruna (UDC), Spain, during October 2013. A group of students of different levels and degrees was targeted. To recruit students from the target groups, we explained to students during class what the experiment would consist of, informed them of the date and time of the sessions, told them they would be invited to a coffee during the tests, and that one of the tests consisted of a game where one of the participants per session would win a prize of 60 euros. In total 126 volunteers, all under- and postgraduate UDC students, participated in the experiment. All sessions took place in a computer room; participants in the same session completed all tests at the same time, each respondent on a separate computer.

Participants completed two types of tests. First, they completed a series of short tests on overconfidence and prospect theory (Peón et al. 2014) to determine their behavioral profile. They then played the strategy game we describe in this paper, aimed at replicating the basics of the decision-making process of a bank granting credit to costumers under conditions of uncertainty and risk. The strategies implemented resulted in three types of indicators (price, quantity and quality of credit) that were tested against the behavioral profile and risk attitudes of the respondents.

In what follows we explain the design of the experimental game. In Sect. 2.1 we briefly review the literature on business simulation games, banking efficiency and credit markets to help us define our approach to game design and to propose a set of dependent variables representative of the credit policies set by participants in the experiment. Those were the dependent variables we tested against the different behavioral profiles of the respondents. In Sect. 2.2 we provide an extensive explanation of the simulation game in the experiment.

2.1 Literature review and basic approach to the experiment

Controlled laboratory experimentation helped economists resolve a major empirical challenge: going beyond correlational analysis to provide insights on causation (List 2009). Furthermore, since experimental economics proved to be a good method for understanding human behavior (Levitt and List 2009), its success is particularly relevant in behavioral economics. Below we provide a brief review of the literature on experimental economics and, in particular, on business simulation games, as the foundation for the type and main characteristics of the research we conducted.

Laboratory and field experiments are used to test a variety of issues, including information assimilation (Levitt and List 2009). Our experiment conducts a simulation of a retail bank for that purpose. A simulation is an evolving case study of a particular social or physical reality in which participants take on bona fide roles with well-defined responsibilities and constraints (De Freitas and Oliver 2006; Gredler 2004). Three necessary elements for an experiment are an environment defining the payoffs, an institution defining language and rules, and the participants' behavior (Smith 2001).

More specifically, Gredler (2004) defines four important requirements for simulations: (1) a model that allows the participants to interact with a complex real-world situation; (2) a defined task and role for each participant involved; (3) an environment that allows participants to execute a range of strategies; and (4) the presence of a feedback system so that participants can change strategies. For our experimental setting, these four characteristics were considered as follows: the simulation game required participants to play the role of a bank granting credit to costumers in a complex situation, where their strategies are defined in terms of credit to be granted, and a feedback system was based on setting a multi-period game where participants could, after each period and before setting a strategy for a new costumer, learn how economic perspectives were evolving and observe their past performance in granting credit.

A classic problem with controlled experimentation is its external validity; specifically, the fact that individuals are in an environment where they are aware that their behavior is being monitored, recorded, and subsequently scrutinized, might cause generalizability to be compromised (Levitt and List 2007). The incorporation of markets, repetition and monetary incentives would improve the validity of the experiment, but perhaps not completely resolve this problem. Consequently, a discussion of the external validity of our experiment is provided in the conclusions.

Additionally, we needed to specify performance indicators that measured the credit policies implemented in the game. Two broad views on what determines how much private credit a financial system would grant are the power theories and information theories of credit (Djankov et al. 2005). Power theories consider that what matters for the viability of private credit is the power of creditors: banks are more willing to extend credit the more easily they can force repayment, grab collateral, or even gain control of the firm (Townsend 1979). For information theories what matters for lending is information: the more banks know about their clients (credit history, financial situation, etc.), the more willing they are to extend credit, since information reduces the 'lemons' problem (Stiglitz and Weiss 1981).

The approach we follow is closer to the informational approach, which, in short, depends on the ability of lenders to screen good borrowers from bad borrowers so that they can implement credit policies that maximize their profits adjusted by risk. The simulation game must provide information about the expected solvency of the potential borrowers. A classic approach to this issue comes from the literature on business failure,[1] in which financial statement analysis plays an essential role. However, if information in the game was given in terms of financial data that participants had to analyze, it would introduce an asymmetry among judges where the more skilled ones would be expected to outperform. Information provided in the experiment had to avoid that, in order to disentangle whether behavioral biases would produce predictable patterns when all participants were given the same, objective information.

Finally, the game indicators were specified in accordance with the literature on banking efficiency. The efficient-structure hypothesis (Demsetz 1973) interprets market power and performance of banks as a consequence of their efficiency levels: banks which operate more efficiently than their competitors will gain higher profits from lower costs, hence they will hold a major market share. Following this hypothesis, different versions of efficiency were examined. A review of the literature on (banking) economic efficiency yielded a list of variables (Berger and Mester 1997) to be used to test cost and profit efficiencies (Table 1), including costs, prices, loan volumes and environmental variables such as non-performing loans (NPL) as a proportion of total loans.

We then proceeded as follows in the experiment. First, the game design had to imitate a complex situation where objective information about macroeconomic perspectives and customers' expected solvency was presented in the form of confidence intervals and updated period by period. The game had to have multiple periods to introduce a feedback system that allowed participants to observe how economic perspectives evolved. Second, the cost structure of the bank was an input, provided to participants as a given variable. Respondents could consequently implement their strategies for each period in terms of prices and volumes of loans granted to a new niche of costumers, given the (updated) information available. Finally, we controlled for loan quality to trace participant risky behavior by

[1] This literature starts with the seminal articles by Beaver (1966), Altman (1968) and (Argenti 1976). See Rodríguez (2000) for a survey.

Table 1 Variables included in the cost and profit functions by Berger and Mester (1997)

	Variables	Description
Dependent variables		
Cost	Variable operating plus interest costs	Includes costs of purchased funds, deposits and labor
Profit	Variable profits	Includes revenues from loans and securities less variable costs
Exogenous variables		
Output quantities	Consumer loans	Including credit cards
	Business loans	All other loans
	Securities	All non-loan financial assets; i.e., gross total assets less (consumer and business loans + physical capital)
Input prices	Price of core deposits	
	Price of purchased funds	
	Price of labor	
Output prices	Price of consumer loans	(Domestic transactions accounts, time and savings)
	Price of business loans	All other liabilities
	Price of securities	
Fixed netput quantities	Physical capital	
	Equity capital	
	Off-balance-sheet items	(Commitments, letters of credit, derivatives) using Basel Accord risk weights to be risk-equivalent to loans
Environmental variables	Ratio of NPL/total loans	NPL = non-performing loans, past due at least 90 days
	Weighted aver. NPL for state/ province	Weighted average using as weight the proportions of the loans issued by banks in the state/province

Source: Berger and Mester (1997)

measuring two ex post variables: their average ratio of non-performing loans to total loans and the ratio between prices granted to high versus low quality consumers.

2.2 Description of the game

The experiment was designed to make participants decide how much credit they would grant to a series of clients and at what price, given certain information about their expected solvency and macroeconomic perspectives. The hypotheses to be tested seek to trace evidence of the effects of different risk profiles of participants on their policy decisions, assessed in terms of prospect theory and overconfidence. We implemented a series of psychological tests (Sect. 3) in accordance with standard tests in the literature. In order to obtain indicators about credit policies we placed special emphasis in several aspects of the game. First, we provided participants with objective information on macroeconomic perspectives and expected solvency, mostly in the form of confidence intervals. Second, we gave them clear instructions regarding how variables in the game were interrelated and how to set their strategies. Finally, to reflect a business cycle that might turn up or down and allow participants' expectations to play a role, we implemented a multi-period game where information was updated after each period.

Each player ran a bank that grants credit to customers. All players played a similar game—facing identical situations of cost and demand and having the same information available to them—but played individually with no interaction with other players. Nonetheless, they did compete, as the objective of the game was to implement the best strategy in terms of profits. To improve external validity, a monetary incentive was introduced: the winner of each session—the participant that earned the highest profit by the end of the game—received a prize of 60 euros. Five rounds of the game were played, with about 20–30 players in each (for a total of 126 participants).

We devised a six-period game in which, at each stage, the bank has access to a different niche set of clients applying for a three-year loan.[2] For simplicity sake, we set the discount rate as equal to zero. Players had to decide how much credit they were willing to grant to that niche and at what price, given the information available about: (a) the niche's expected default rates in the form of confidence intervals[3]; (b) macroeconomic perspectives regarding GDP growth and Euribor-_1Y rates, also in the form of confidence intervals; (c) and calculations of (ex ante) expected profits and delinquency ratios given the player's strategy, as well as the ex post profits and default rates obtained in each period after their strategies had been set and the information had been updated. Figure 1, which is a screenshot of the computer application for the game in period 1, shows information provided on macroeconomic perspectives (above), niche default rates and player strategies (left) and expected and historical profits and portfolio delinquency ratios (right).

Macroeconomic data. At the beginning of each stage participants were given information about economic perspectives (expected GDP growth and Euribor_1Y rates) in terms of confidence intervals. Both numerical and graphical information are shown for periods 1–8.[4] Graphics allow for a more intuitive interpretation, in particular when information is to be updated in subsequent periods: confidence intervals use different colors, a thin line represents the last year's estimates, a dotted line the initial (period 1) estimates, and shadowed areas represent past periods (see Fig. 2 referring to period 3).

[2] Note that niche clients were different at different periods, but all participants faced an identical niche in the same period. We considered six periods to be enough for the two purposes we implemented a multi-period game: to set a feedback system that allows participants to learn how economic perspectives evolved, and to have a larger data set of strategies implemented by them. Finally, we considered potential borrowers to be applying for a three-year loan in order to make decision-making more complex; thus, in order to grant credit to a niche, participants had to consider the possible economic scenarios for the next three periods, not just one.

[3] Confidence intervals were given in three-point format: average values and high and low boundaries. Boundaries were explicitly said to be absolute limits that could not be surpassed. That meant, for instance, that if the expected default rate of a niche is (15, 5, 1 %), the highest (lowest) default rate in all possible states of the world is 15 % (1 %). That also meant, for instance, that if in period 1 we said that expected GDP growth for period 5 could range between (-1, 1, 3 %), updated information in periods 2–4 could not say that the expected GDP growth for period 5 might go higher than 3 % or lower than -1 %.

[4] It is a six-period game, but in every period a three-year loan is granted.

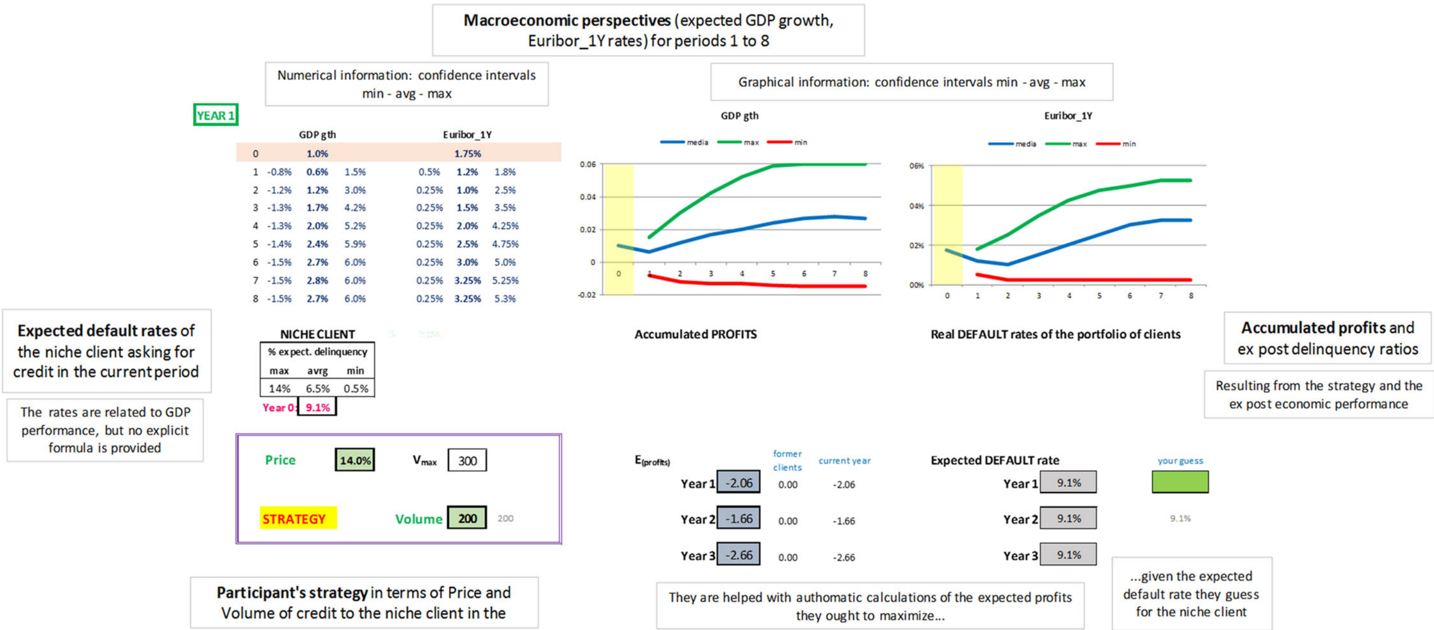


Fig. 1 Screenshot of the game at period 1

YEAR 3

		GDP gth				Euribor_1Y		
0		**1.0%**				**1.8%**		
1		**-0.1%**				**1.0%**		
2		**0.5%**		0.9%		**0.4%**		-0.1%
3	0.1%	**1.3%**	2.8%	0.3%	0.4%	**0.8%**	2.0%	-0.3%
4	0.3%	**2.1%**	4.4%	0.0%	0.5%	**1.3%**	3.0%	-0.5%
5	0.3%	**2.8%**	5.5%	0.1%	0.6%	**2.0%**	4.1%	-0.3%
6	0.2%	**3.2%**	6.0%	0.2%	0.5%	**2.8%**	4.8%	-0.2%
7	0.0%	**3.5%**	6.0%	0.3%	0.5%	**3.5%**	5.2%	0.1%
8	-0.2%	**3.6%**	6.0%	0.4%	0.4%	**3.8%**	5.3%	0.2%

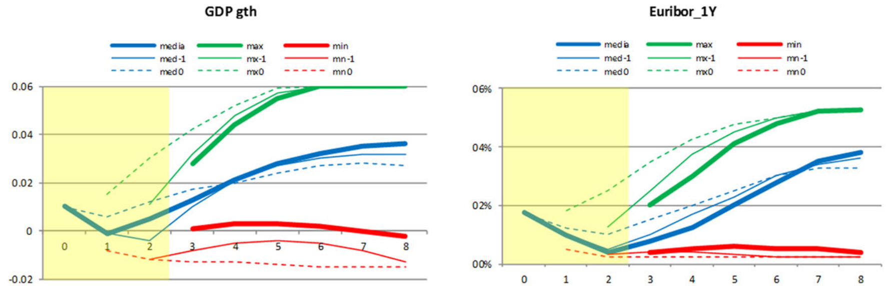


Fig. 2 Updated macro information at period 3 (color figure online)

Numerical data also shows how economic perspectives changed: positive and negative variations in mean estimates for GDP growth and Euribor rates between two consecutive periods are highlighted in blue and red, respectively (see Fig. 2).

Information not given to participants. GDP growth rates were designed to range, at most, from -1.5 to 6.0 % with an average (mean economic performance) of 2.5 %. Boundaries were wider the farther the estimate from the current period. On average, the amplitude of the intervals would be about 1 % for an estimate for the next year, 1.75 % 2 years ahead, 2.5 % 3 years ahead, and so on—though actual ranges may vary to some extent. Euribor rates were designed to vary, intuitively, in accordance with GDP perspectives, with lower (higher) rates being correlated with weak (stronger) GDP perspectives.

Since participants first played a version of the game for practice, and then played a second version of the game where they would compete for the prize, we devised two alternative scenarios, denoted practice version and game, respectively. These are summarized by the ex post economic data at the end of the game in period 8 (see Fig. 3).

We designed a scenario for the game where the macroeconomic perspectives tended to improve—as would happen during an upturn in the economic cycle. The GDP and Euribor values were not tied to any specific country nor they were related to real-world values. Participants were not told whether they should expect this scenario to be realistic or not.

Niche default rates. Information about expected default rates for the six niche clients at each stage was given in terms of confidence intervals (Table 2).

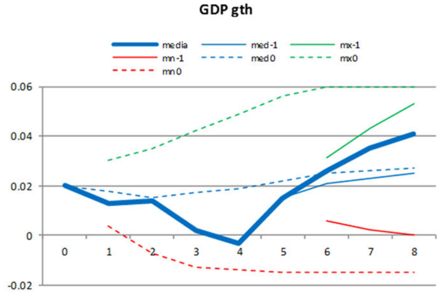
Players were only told that maximum, minimum and average default rates were associated with the weakest, strongest and average GDP performance rates, respectively, but received no information about the explicit mathematical relationship between GDP growth and delinquency. Players were told, when setting their price and credit volume strategies, to infer the expected default rate of the niche client given the information provided for the niche and for macroeconomic perspectives. As a starting clue in period 1, explicit information about the true default rate of the niche in the previous period (period 0, before the game started) was provided. For the subsequent periods no such information was given, since players could learn the ex post default rates for their portfolios once economic scenarios were updated.

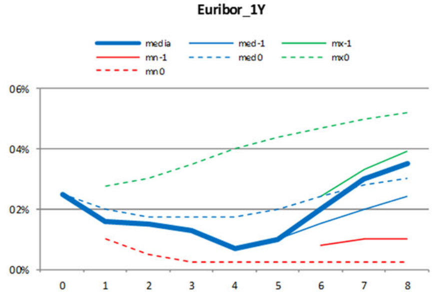
Information not given to participants. The six niche clients are of two types, according to their expected default rates. Type A clients are riskier than type B, both because they exhibit wider intervals and because downside risk is substantially higher. Ex post default rates were computed given GDP performance. In particular, reference GDP growth rates were estimated as 2/3 the current year's (ex post) rate plus 1/3 the previous year's rate. Then, the default rates were set to fall at the equivalent point in the interval as the reference GDP rate within the (-1.5, 2.5, 6.0 %) interval mentioned above. For instance, for niche C1 if the ex post GDP growth rate in period 1 was -0.1 % and the previous year's rate was 1.0 % we have:

Practice version

7 & 8

	GDP gth		Euribor_1Y	
0	2.0%		2.5%	
1	1.3%		1.6%	
2	1.4%		1.5%	
3	0.2%		1.3%	
4	-0.3%		0.7%	
5	1.5%		1.0%	
6	2.6%	0.5%	2.0%	0.5%
7	3.5%	1.2%	3.0%	1.0%
8	4.1%	1.6%	3.5%	1.1%

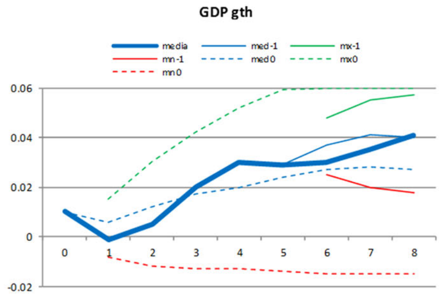




Game

7 & 8

	GDP gth		Euribor_1Y	
0	1.0%		1.8%	
1	-0.1%		1.0%	
2	0.5%		0.4%	
3	2.0%		0.8%	
4	3.0%		2.3%	
5	2.9%		3.2%	
6	3.0%	-0.7%	3.8%	1.0%
7	3.5%	-0.6%	3.0%	-0.5%
8	4.1%	0.1%	2.5%	-1.3%



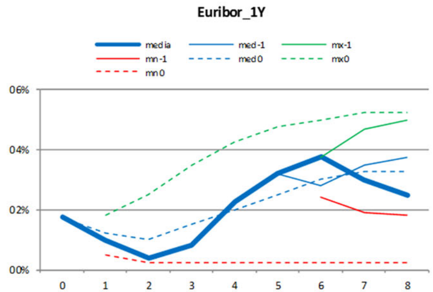


Fig. 3 Scenarios of the practice version and game at period 8

Table 2 Expected default rates of all niche clients

Type	Niche	% Expected default rate		
		Max (%)	Mean (%)	Min (%)
A	C1	14	6.5	.5
B	C2	10	5	2
B	C3	9	4	1
A	C4	16	7	0
A	C5	18	7.5	1
B	C6	11	5	1

C1–C6 denote the six niche clients of two different types, A (riskier) and B

$$Reference\ GDP\ rate = \frac{2}{3}(-0.1\,\%) + \frac{1}{3}(1.0\,\%) = 0.267\,\%. \tag{1}$$

Since 0.267 % lies in the bottom half of the interval (−1.5, 2.5 %), we set the ex post default rate for period 1 to lie at the (linearly) equivalent point within the bottom range (14, 6.5 %)—see data in Table 2—which is equivalent to computing

$$Real\ default\ rate = 14\,\% + ratio(0.267 - 2.5\,\%), \tag{2}$$

where $ratio = (14 - 6.5\ \%)/[2.5\ \% - (-1.5\ \%)] = -\ 1.88$ measures the impact of a one percent increase in GDP on the reduction of the default rate. In our example for client C1, the ex post default rate in period 1 would be 10.69 %.

Strategy. In every period, players had to analyze the information available and determine their strategy. Strategies were defined in terms of two variables: the price, p, at which they are willing to grant credit to that niche of clients, and the volume of credit granted to that niche. For this purpose, they were helped with automatic calculations—right hand side of the screenshot in Fig. 1—of the expected profits they ought to maximize.

Participants were instructed to proceed to make an initial guess of the expected default rate for that niche market ('your guess' cell in Fig. 1) and then to set a price that ranges within 10.0 and 20.0 %. For their price, they were given the maximum volume demanded by the niche market, V_{max}, which reflects a linear demand function[5] for credit as

$$V_{max} = 1000 - 5000p, \tag{3}$$

hence credit volumes ranged from 0 to 500 euros. Given V_{max}, they had to decide whether they grant the maximum volume of credit demanded, that is, $V = V_{max}$, or they set V such that $0 \leq V < V_{max}$.[6] This allowed us to measure which participants were being more conservative in terms of volume—rather than just having price as the

[5] The demand function was not provided to players, but they were obviously given the outcome V_{max}.

[6] More specifically, they were instructed as follows: "Please note the computer application helps you to calculate the expected profits given the inputs being set ($E[m]$, p, V). Be aware that these are the inputs that you set; the expected profits may be fulfilled or not depending on whether (a) the economic scenario follows the path you anticipated; and (b) the strategy you consequently implemented is indeed optimal. Therefore, be advised, when setting your strategies, that the expected profits are just an aid. On one hand, not granting credit, $V = 0$, when you think a niche may not generate profits allows you to save a fixed cost of 3 euros. On the other hand, if you decide to give credit, granting $V = V_{max}$ or a lower volume should depend on how sure you are this niche client is going to render you profits rather than losses".

single decision variable. Hence, we observed two variables per participant, p and V, that might not clear the market (whenever $V < V_{max}$). In the instructions players were advised to act depending on "how sure you are your strategy is going to be profitable", since setting $V = 0$ also made fixed costs equal to zero (see next paragraphs).[7]

Profit calculations. To help participants set their strategies, the computer application provided ex ante estimates of profits and portfolio default rates. Players were also given information, before the game started, regarding the mathematical expression for the profit function (income minus costs) used to derive those values, although they were told that this was merely given so they could better understand the game and were advised to follow the calculations instead. The (expected) revenue function was set as:

$$E[R] = Vp(1 - E[m]) - VE[m], \tag{4}$$

where $E[m]$ is the expected default rate of the niche (provided by the player or estimated by default otherwise).[8] The cost function is

$$C = F + (E[e] + v)V, \tag{5}$$

which includes a fixed cost, F, of 3 euros per period and the number of active niche clients (i.e., those for which $V \neq 0$ in the last 3 years), a variable cost (Euribor, e) equal to the (expected) Euribor-1Y rate provided (representing the cost of the deposits needed to fund granted credit), and another variable cost, v, of 2.0 %, as the cost of managing a higher volume of credit.

From (4) and (5) the expected profit function is

$$E[\pi] = \{(1 - E[m])p - (E[m] + E[e] + 2.0\,\%)\}V - 3, \tag{6}$$

whereas the ex post profit observed amounts to

$$\pi = \{(1 - m)p - (m + e + 2.0\,\%)\}V - 3, \tag{7}$$

where m and e replace the ex ante expectations $E[m]$ and $E[e]$, respectively.

Finally, players were also given information about the delinquency ratios of their portfolios (right hand side of the screen; see Fig. 1). Regarding expected delinquency, players could make their own estimates ('your guess' cell in Fig. 1). If that cell was left blank, the computer set a default expected default rate: in period 1, the ex post default rate of the niche in period 0 and, in subsequent periods, the

[7] For indicator estimates (see "Game Indicators" in this Sect. 2.2), when a subject sets $V = 0$ we set $p = 20\,\%$, i.e., the price that should be offered to have zero demand, disregarding the actual value the participant set. We did so in order to have indicators that were homogeneous across participants: judges set $V = 0$ after they tried different prices (sometimes even providing no answer for p), so the last price they set may not be representative. This correction did not apply to any other case since, as explained, we wanted to observe both price and volume strategies that might not clear the market.

[8] That income function makes two implicit assumptions. First, a default means the bank recuperates neither interest nor capital from that proportion m_e of the loan. Second, for simplicity sake and easier interpretation by the players, we assumed the total credit granted to a niche in all 3 years the credit was active to be equal to the initial V granted. That may be interpreted as a line of credit to a niche of clients that is renewed annually for the total amount, independent of the default rate incurred in any previous year(s). Participants were explicitly informed of both assumptions.

Table 3 Optimal ex post strategies for all niche clients

Type	Niche	Observed default rates (%)									Optimal strategy	
		m1	m2	m3	m4	m5	m6	m7	m8	Average	*p**	V*
A	C1	10.7	10.6	8.4						9.9	17.0 %	*149.6*
B	C2		7.8	6.3	4.9					6.3	15.0 %	*248.3*
B	C3			5.3	3.9	3.6				4.2	14.3 %	*282.6*
A	C4				6.7	6.1	6.1			6.3	16.1 %	*197.1*
A	C5					6.7	6.6	6.0		6.4	16.3 %	*186.2*
B	C6						4.5	4.0	3.4	4.0	14.7 %	*264.3*
										Type A	**16.4 %**	*16.4%*
										Type B	**14.7 %**	*14.7%*
										Price ratio	**1.119**	**1.117**

observed default rate for his or her current portfolio of clients. The historic ex post default rates were also provided to the player, computed as the weighted average (weighted by V) of the observed default rates for all active niche clients in the player's portfolio.

Optimal ex post strategies. We may compare participants strategies with what would be the optimal *ex post* strategies (i.e., once we know in period 8 how the economy actually performed). We compute these by setting $V = 1000 - 5000\,p$ in Eq. (7) and then deriving with respect to p, as follows,

$$p^* = \frac{1000 + (5000 - 1000) \cdot \bar{m} + 5000 \cdot (\bar{e} + 2.0\,\%)}{2 \cdot 5000 \cdot (1 - \bar{m})} \tag{8}$$

where $\bar{m}$ and $\bar{e}$ are the average values of the default and Euribor rates, respectively, during the three years of the loan to a particular niche client. Table 3 summarizes the optimal ex post strategies for all six niches in the game.

These computations were helpful to confirm that type A clients, considering their risky profile, should be charged a higher price, but also taking macroeconomic performance into account. The average optimal price for type A niches was higher than for type B [16.4 vs. 14.7 % in both average (bold) and volume-weighted average (italics) data], as expected. The price ratio suggests that players should set a price to type A clients that is a 12 % higher (in relative terms, for a price ratio of about 1.12) than for type B clients. This information was used as a benchmark for a quality indicator later on (see next section).

Game indicators. At the end of the game, each player yielded 6 × 2 decision variables: a pair (p, V) that represented the credit they were willing to grant to each of the six niche clients. As mentioned above, based on this information we wanted to trace differences between players regarding three types of indicators:

- *Price indicators* We computed two estimators to account for differences in price strategies: p_{avg}, the average price across the six niches; and p_{vol}, the volume-averaged price across the six niches. We compute the volume-averaged indicator because two participants that set the same price for a given niche may not grant the same volume of credit, as we allowed $V < V_{\max}$.

- *Volume indicators* We computed two volume estimators, denoted for simplicity sake as V_{ind} and $V_{\max,\text{ind}}$. V_{ind} sums the volumes granted to each of the six niches (i.e., $V_{\text{ind}} = \Sigma V$), whereas $V_{\max,\text{ind}}$ compares the volumes actually granted by the player with the demand available for the price she set. We computed $V_{\max,\text{ind}}$ as the ratio between V_{ind} and the sum of all $V_{\max}$ that would be demanded by each niche, given the prices the respondent had set.
- *Quality indicators* We measured the risky behavior of participants by computing two types of indicators. First, the NPL indicator measures the average ratio of non-performing-loans to total loans. Second, we compared the prices granted to clients of low (type A) versus high (type B) qualities. Likewise we did with price indicators, we computed two different indicators for this purpose: Q_{avg} was calculated as the ratio of the average price to type A clients over the average price to type B clients, while Q_{vol} sets a similar ratio for volume-weighted average prices. Table 3 offers a benchmark for these quality indicators: the optimal ex post strategies entail the ratios 1.119 and 1.117 for mean and volume-average data, respectively. Hence, when indicators Q_{avg} and Q_{vol} for a given player are well below those levels, it indicates an aggressive pricing strategy in favor of low-quality borrowers.

Table 4 summarizes all these seven indicators, which were the dependent variables in the hypothesis testing process (Sect. 3).

Game implementation. Participants were instructed how to play the game in three steps. Firstly, they were given an extensive explanation of the game. Secondly, after being provided with a set of written instructions that summarized all the rules, participants had time to play a version of the game for practice. Finally, once participants confirmed they understood the game and had clarified doubts, they played the game, with players earning the maximum profits winning cash prizes. The results obtained in this second version of the game were the only used for research purposes.

Design limitations. Despite our efforts to devise an experimental setting that mirrors real loan decisions, it does not exactly correspond to the more complex nature of banks and their institutional environment. Some limitations are discussed below in terms of the extent to which we believe they might compromise the generalization of the experiment results.

A first limitation of the analysis is the evidence that internal processes for granting credit by banks are more complex than decisions made by a single individual. This would involve different departments (loan offices versus back offices) and specific regulation under the Basel Accords. Being this true, the process of granting credit may be simplified to two main steps: bank executives set the basic credit policies according to the cost of funding and expected macroeconomic performance, and then pass these policies down to employees and risk analysis departments in the form of commercial goals. The literature on behavioral corporate finance has shown that the performance of firms is largely affected by the behavioral profile of executives in the firm; take, for instance, the literature on business failure

Table 4 Summary of game indicators

Variable	Values	Interpretation	Calculation	Literature
p_{avg}	$p^* \rightarrow$ min 10.0 % – max 20.0 %	↓ Price → ↑ risk strategy	Average price across 6 niche clients	Defining relevant indicators of the game: Berger and Mester (1997)
p_{vol}			Volume-weighted average price across 6 niches	
V_{ind}	$V_{ind}^* \rightarrow$ min 0 – max 500	↑ Volume → ↑ risk strategy	Average volume of credit granted (6 niches)	
$V_{max,ind}$	$V_{max,ind} \leq 1$ where $V_{max,ind} = 1 \rightarrow$ full credit at p^*		$V_{max,ind} = V_{ind}/(\sum_{6\ niches}[V_{max}\vert p^*])$	
NPL	% of non-performing loans (min 0 %)	↑ NPL → ↑ risk strategy	Average *ex post* NPL ratio across 6 niche clients	Design and measurement of indicators: own elaboration
Q_{avg}	$Q_{avg} < 1 \rightarrow$ lower p to risky niches $(Opt._{expost}\ Q_{avg} = 1.119)$	↓ Quality → ↑ risk strategy	Mean prices to costumers of high versus low qualities	
Q_{vol}	Idem $(Opt._{expost}\ Q_{vol} = 1.117)$		Idem, volume-weighted	

and CEO overconfidence (cited above). Hence, we can assert that the decisions of a single, or a few, decision makers do indeed determine, to some extent, the credit policies of banks. The literature on social contagion (Asch 1952; Shiller 1984), obedience to authority (Milgram 1963, 1974), and groupthink theory (Lunenburg 2010), moreover, explain how these biases could propagate inside firms. Thus, the commercial success of loan officers is often used as feedback by executives when re-evaluating their economic expectations, in such a way that a trend might be generated by the feedback effect of the commercial goals achieved (Peón and Calvo 2012). However, these effects of social contagion are much too complex to be analyzed in an experimental setting.

A second drawback is that the game does not account for aspects of the institutional environment such as capital regulation. For simplicity sake, we have not considered the effects of participants' decisions on equity performance (basically, we obviate default effects or assume banks have access to an unlimited source of capital). This reduces the complexity of the game, but introduces an asymmetry in the judges' evaluation—since they do not pay attention to the effects of accumulated losses on future bank solvency. We consider that this effect by itself should not have a relevant impact on the participants' strategies. However, the effect might not be negligible if we take into account the third limitation.

The third and more important limitation is the evidence that the incorporation of monetary incentives, though often desirable to improve the external validity of field experiments, might induce participants to play strategically. Players might choose to play too aggressively if they bet on future profits or too conservatively if they bet on future losses, only to deviate from what they expect other respondents will play.

That is, some participants might choose extreme policies merely to play differently from their competitors—particularly if they are behind in the game. The external validity of the results will be discussed in the conclusions, but for now we anticipate four comments on this debate. Firstly, the monetary incentives introduced are consistent with how executives at banks are usually paid: according to their ability to outperform competitors. Second, the combination of incentives and the evidence that the negative effects of their decisions will be borne by others would induce a moral hazard behavior that is often alleged to be a reason behind the recent worldwide crisis (e.g., Acharya and Naqvi 2012). This moral hazard effect may be compared in our experiment to the combination of monetary incentives and the absence of considerations in regards to capital solvency. Third, by introducing indicators that measured player performance in terms of a variable they are not aware of—the quality of the niche markets—we have a powerful tool to analyze the external validity of the results, regardless of the strategic behavior of participants in terms of price and volume of credit. Finally, at the very least, if the game design induced risk seeking behavior, it would be of interest to assess which of the different manifestations of overconfidence and prospect theory might foster such aggressive behavior.

To sum up, in spite of these limitations, our experimental game may be considered to synthesize the main stylized features of real lending decisions.

3 Tests

The goal of the experiment is to trace the effects that different behavioral profiles might have on credit policies implemented by different participants. To such purpose, before the players entered the strategy game, a series of tests were implemented to determine their psychological profiles in terms of overconfidence and prospect theory. Below we describe the behavioral tests (Sect. 3.1) and the hypotheses to be tested (Sect. 3.2). We leave the analysis of the results for a separate section.

3.1 Behavioral tests

We designed a series of tests based on standards in the literature on overconfidence and prospect theory. Thus, on one hand we follow Moore and Healy's (2008) theory on the three different measures of overconfidence, and design shorter versions of Soll and Klayman's (2004) and Moore and Healy's (2008) tests to elicit those measures at the individual level. On the other hand, in regards to prospect theory, we follow Rieger and Wang's (2008) normalization of prospect theory (Kahneman and Tversky 1979) assuming classic parametric functions in the literature, while for test design we merge some features of Tversky and Kahneman's (1992) elicitation method merged with the approach for making an efficient test with a minimum number of questions by Abdellaoui et al. (2008). A brief description of the tests is provided below, along with an explanation of how the key psychological constructs were measured. For an extended interpretation and analysis of the validity of the results obtained with our tests, see Peón et al. (2014).

Overconfidence. We calculated three different measures: overestimation, **E**, which measures the degree to which a respondent overestimates her own performance; overplacement, **P**, which measures the degree to which she overestimates her performance relative to others; and overprecision, **M**, which measures an excessive precision to estimate future uncertainty (miscalibration). Our test for **E** and **P** is a simple version of Moore and Healy's (2008) trivia tests: a set of four quizzes with 10 items each, involving general knowledge questions with a time limit to answer two quizzes of easy and two of hard difficulty to account for the hard-easy effect. After time was up for each round, participants were required to estimate their own scores and the average score of the other students participating in the experiment.

E was computed by substracting a participant's actual score in each of the four trivia, x_i, from her reported expected score, $E[X_i]$, and then summing all four results. A measure $\mathbf{E} > 0$ implies overestimation and $\mathbf{E} < 0$ means underestimation. Overplacement takes into account whether a participant is really better than others. For each quiz we use the expression

$$\mathbf{P} = \left(E[X_i] - E\left[X_j\right]\right) - \left(x_i - x_j\right) \tag{9}$$

where $E[X_i]$ is that person's belief about the expected average score on that quiz of the other participants and x_j measures the actual average score, and then sum all four results. A measure $\mathbf{P} > 0$ implies overplacement, while $\mathbf{P} < 0$ means underplacement.

Overprecision was analyzed through a separate set of six questions on confidence interval estimates. To disentangle true overprecision from variability in setting interval widths we follow Soll and Klayman (2004) in computing the ratio $\mathbf{M}$ = MEAD/MAD, where MEAD is the mean of the expected absolute deviations implied by each pair of fractiles for a subject, and where MAD is the observed mean absolute deviation. **M** represents the ratio of observed average interval width to the well-calibrated zero-variability interval width. Thus, $\mathbf{M} = 1$ implies perfect calibration, and $\mathbf{M} < 1$ indicates an overconfidence bias that cannot be attributed to random error.

Following Soll and Klayman (2004) we devised our test as follows. Participants had to specify a three-point estimate (median, 10 and 90 % fractiles) for three pairs of questions in three different domains—two traditional almanac questions (the year a device was invented and mortality rates), plus a domain for which participants could draw on direct, personal experience (the time required to walk from one place to another). In order to estimate **M**, we use a beta function to obtain the implicit subjective probability density function for each respondent, then we estimate MEAD and MAD for each pair of questions per domain and, consequently, a ratio M for each domain. **M** could then simply be calculated as the average ($\mathbf{M}_{\mathrm{avg}}$) or the median ($\mathbf{M}_{\mathrm{med}}$) of the three different estimations (one per domain).

Prospect theory. We elicited participant preferences following Rieger and Wang's (2008) normalization of prospect theory (Kahneman and Tversky 1979) to obtain value and weighting functions assuming two classic parametric specifications. First, the piecewise power value function by Tversky and Kahneman (1992),

$$v(x) = \begin{cases} x^{\alpha^+}, & for\, x \geq 0 \\ -\beta(-x)^{\alpha^-}, & for\, x < 0 \end{cases} \tag{10}$$

where x accounts for gains (if $x \geq 0$) or losses (if $x < 0$), α^+ measures the sensitivity to gains, α^- the sensitivity to losses, and β measures loss aversion. Second, the Prelec-I weighting function (Prelec 1998) given by

$$w(p) = \exp(-(-\log(p))^{\gamma}) \tag{11}$$

where $\gamma > 0$, to estimate the probability weighting function, with decision weights $w(p)$ subsequently normalized so they add up to one. The obtained parameters allowed us to determine, for each participant, the curvature of the value function for gains and losses (α^+ and α^-), the degree of loss aversion (β), and the distortion of probabilities for gains and losses (γ^+ and γ^-) in the weighting function.

Although cumulative prospect theory (CPT) is more frequently cited in the literature, Rieger and Wang (2008) observe that not all properties of CPT correspond well with experimental data. Normalized prospect theory (NPT) has some advantages over CPT (it cures the violations of state-dominance in lotteries with two outcomes and its utility converges to a continuous distribution), while it is also an easier approach to compute which, in particular, simplifies the computation of the loss aversion parameter in our questionnaires. For such purpose, NPT has been recommended for eliciting the preferences of a given individual in a simple manner (e.g., Hens and Bachmann 2008, in the context of private banking).

For parameter estimation, our method merges some characteristics of Tversky and Kahneman's (1992) approach to elicit certainty equivalents of prospects with just two outcomes and Abdellaoui et al. (2008)'s proposal to make an efficient test with a minimum number of questions. Thus, the elicitation method consists of three stages, with 15 questions in total: six questions involving only positive prospects (i.e., a chance to win some positive quantity or zero) to calibrate α^+ and γ^+; six questions for negative prospects to calibrate α^- and γ^- (using a nonlinear regression procedure separately for each subject); and finally, three questions regarding the acceptability of mixed prospects in order to estimate β. Prospects devised to calibrate α^+, γ^+, α^-, and γ^- used significant, albeit hypothetical, sums of money: 500, 1000 and 2000 euros. The three questions to estimate loss aversion used smaller amounts: utility is close to linear for moderate amounts of money (Rabin 2000), what allows us to assume $\alpha^+ = \alpha^- = 1$ to simplify the estimation of β (as either a mean or median across prospects, denoted β_{avg} and β_{med}, respectively).

For questions with only positive prospects, in each iteration participants had to choose between a positive prospect and a series of positive, sure outcomes. Every time a subject chose either the prospect or the sure gain, a new outcome was provided. This process was repeated until the question was completed and the player could continue with another prospect. The probabilities of success were different (two questions with probability of success of 50 % and one with probabilities of 99, 95, 5 and 1 %); this information was emphasized to help avoid wrong answers. Following Abdellaoui et al. (2008), to control for response errors we repeated the last sure outcome of the first series at the end of each trial. The certainty equivalent of a prospect was then estimated by the midpoint between the lowest accepted value and the highest rejected value in the second set of choices. Tversky and Kahneman (1992) indicate that this procedure allows for the cash equivalent to be derived from observed choices, rather than assessed by the subject. Finally, the procedure was similar for questions involving only negative prospects, except that now prospects and sure outcomes were expressed in terms of losses and probabilities were expressed in terms of probabilities of losing. Certainty equivalents were estimated similarly (for values in absolute terms).

Table 5 summarizes the results for the overconfidence and prospect theory tests for our experiment. Note that, given the way the variables were defined in our tests, overconfidence increases the higher **E** and **P** and the lower **M**, risk seeking increases the higher α^{+} and the lower α^{-}, loss aversion increases with β and distortion of probabilities increase the lower γ^{+} and γ^{-}.

Peón et al. (2014) analyze the validity of the results for this research. In brief, the trivia tests allowed us to replicate the standard results for individual measures of overestimation and overplacement. In addition, our test for risk attitudes provided similar results to those observable in the theoretical and empirical literature on prospect theory, considering the properties of the value and weighting functions, the fourfold pattern of risk attitudes, the iteration and fitting errors measured for each participant, and anomalies detected at the individual level. However, our test on interval estimates for overprecision provided individual estimations that vary if different refinement methods are used—with differences being more accused if average estimates across domains are computed. For future research, we recommend including more questions per domain, and also asking additional questions on personal experience to balance domains. However, for the purposes of hypothesis testing in this article, the robustness of the effects of overprecision were determined by comparing results for the miscalibration ratio **M**—a valid option, as it uses all information provided by the respondent—with results obtained under the other two refinement methods.[9]

Finally, however we have chosen standard tests and specifications in the literature to perform our tests, we must be aware that different ways of measuring biases may yield different results. Hence, the robustness of the results we obtain in the hypothesis testing depends on the ability to replicate similar results under different elicitation methods.

[9] The original refinement of **M**, already described, takes estimates of MEAD and MAD based on the beta function that best fits the three point estimates by the respondent. Alternatively, Soll and Klayman (2004) suggest measuring MAD by assuming the median is in the middle of the distribution, denoted M_2. A third measure is where both MEAD and MAD computations assume a normal distribution, denoted M_N. Only median estimations of these two alternatives were considered since, given the nature of the reliability problem observed in our test, average estimates were shown to be less reliable than medians.

Table 5 Descriptive statistics of behavioral biases

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Age	126	36.00	17.00	53.00	22.15	3.72	13.825	4.704	.216	37.433	.428
Level	126	6.00	1.00	7.00	4.04	2.22	4.918	.155	.216	−1.400	.428
E	126	28.00	−8.00	20.00	2.93	4.76	22.643	.790	.216	1.529	.428
P	126	27.00	−13.98	13.02	−2.71	4.69	21.959	.302	.216	.785	.428
M_{med}	125	1.50	.00	1.50	.34	.26	.066	1.841	.217	4.902	.430
M_{avg}	125	1.32	.07	1.38	.46	.29	.085	1.310	.217	1.837	.430
Alpha+	126	2.43	.24	2.67	1.02	.46	.213	1.513	.216	2.482	.428
Alpha−	126	2.24	.05	2.29	.52	.31	.098	2.320	.216	9.199	.428
Gamma+	126	.95	.05	1.00	.64	.26	.065	−.163	.216	−.700	.428
Gamma−	126	.95	.05	1.00	.53	.28	.077	.183	.216	−1.147	.428
β_{med}	126	9.40	.60	10.00	3.01	1.97	3.897	1.599	.216	3.182	.428
β_{avg}	126	26.00	.67	26.67	3.64	3.57	12.750	3.978	.216	20.157	.428

3.2 Hypothesis testing

To test whether more aggressive profiles (in terms of behavioral biases) were correlated with riskier credit strategies in the game, we define a risky credit strategy as one that is based on lower prices (p_{avg} and p_{vol}) and higher volumes (V_{ind} and $V_{max,ind}$), and yields more non-performing loans (higher NPL ratios) and lower quality ratios (Q_{avg} and Q_{vol}), since these ratios were defined in terms of low over high quality clients. We also define an aggressive profile as having at least one of the following features: a risk-seeking profile, lower loss aversion, and higher overconfidence. Lower loss aversion and higher overconfidence were implied by lower β and **M** values and higher **E** and **P** values. However, a risk-seeking profile could manifest itself through more complex instances.[10]

It holds that the higher α^+ and the lower α^-, the greater the risk seeking (ceteris paribus for similar probability weights). Similarly, lower γ^+ implies more risk seeking but for low probability gains only, while lower γ^- implies more risk seeking but only for moderate/high probability losses. A key assumption henceforth will be moderate/high probabilities for gains and low probabilities for losses, as relevant in terms of probability weighting: player strategies depended on expected delinquency ratio, but we set the highest probability of default for all six niches to be below 20 %. Hence, we define risk seeking behavior in terms of higher γ^+ (for high probability gains) and higher γ^- (for low probability losses).

We test whether more aggressive profiles (risk seeking, loss aversion, and overconfidence) were related to more risky credit strategies. This implies a set of hypotheses as follows.

1. To test whether risk seeking behavior had a predictable effect on credit policies we pose the following hypotheses: the more risk seeking (1a) the lower the price charged to clients, (1b) the higher the volume granted, (1c) the higher the ratio of non-performing loans to total loans, and (1d) the lower the quality ratios.
2. To test whether loss aversion had a predictable effect on credit policies we pose the following hypotheses: the lower the loss aversion (2a) the lower the price charged to clients, (2b) the higher the volume granted, (2c) the higher the NPL ratio, and (2d) the lower the quality ratios.
3. Finally, to test whether any of the three measures of overconfidence explained riskier behavior in the credit market we pose the following hypotheses: the more overconfidence (3a) the lower the price charged to clients, (3b) the higher the volume granted, (3c) the higher the NPL ratio, and (3d) the lower the quality ratios.

[10] The fourfold pattern of preferences in prospect theory implies that risk aversion depends on curvature of the value function and probability weighting simultaneously. Additionally, given the inverse S-shape of the weighting function, a given probability distortion implies different risk profiles for low and medium/high probabilities. Consequently, in this paper, whenever we make a statement like “the higher α^+ the more risk seeking” we are ignoring the effect of probability weightings, and the reverse.

Table 6 Descriptive statistics of game indicators

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
p_{avg}	126	5.67 %	12.75 %	18.42 %	16.05 %	.0104	1.082	−.306	.216	.083	.428
p_{vol}	126	5.49 %	12.19 %	17.68 %	15.40 %	.0109	1.190	−.457	.216	.029	.428
V_{ind}	126	1459	371	1830	1103	269.10	72417.2	−.013	.216	.243	.428
$V_{max,ind}$	126	.78	.22	1.00	.94	.10	.011	−3.721	.216	19.528	.428
NPL	126	2.00 %	5.11 %	7.11 %	5.95 %	.0033	.111	.144	.216	1.021	.428
Q_{avg}	126	.70	.69	1.39	1.02	.11	.012	.491	.216	1.417	.428
Q_{vol}	126	0.69	.74	1.42	1.02	.10	.010	.390	.216	2.083	.428

Table 7 Correlations among behavioral variables

	E	*p*	M_{med}	M_{avg}	Alpha+	Alpha−	Gamma+	Gamma−	β_{med}	β_{avg} (r)
E										
Pearson correlation	1	***.690*****	−.123	*−.166*	−.055	*.165*	.055	−.025	−.065	***−.199****
Sig. (2-tailed)		***.000***	.173	*.064*	.544	*.065*	.537	.782	.468	***.027***
N	126	126	125	125	126	126	126	126	126	123
p										
Pearson correlation	.690**	1	−.039	−.144	−.122	.096	.104	.069	−.010	−.089
Sig. (2-tailed)	.000		.664	.109	.174	.285	.246	.441	.913	.325
N	126	126	125	125	126	126	126	126	126	123
$\mathbf{M_{med}}$										
Pearson correlation	−.123	−.039	1	.672**	−.121	.008	.032	*.165*	.052	.087
Sig. (2-tailed)	.173	.664		**.000**	.180	.932	.723	*.066*	.564	.339
N	125	125	125	125	125	125	125	125	125	122
$\mathbf{M_{avg}}$										
Pearson correlation	−.166	−.144	.672**	1	−.095	.041	−.021	.144	.122	***.193****
Sig. (2-tailed)	.064	.109	.000		.293	.653	.812	.110	.175	***.033***
N	125	125	125	125	125	125	125	125	125	122
Alpha+										
Pearson correlation	−.055	−.122	−.121	−.095	1	***−.211****	***.597*****	−.133	−.036	−.040
Sig. (2-tailed)	.544	.174	.180	.293		***.018***	***.000***	.139	.687	.658
N	126	126	125	125	126	126	126	126	126	123
Alpha−										
Pearson correlation	.165	.096	.008	.041	−.211*	1	−.093	***.326*****	***.294*****	***.308*****
Sig. (2-tailed)	.065	.285	.932	.653	.018		.301	***.000***	***.001***	***.001***
N	126	126	125	125	126	126	126	126	126	123

Table 7 continued

	E	p	M_{med}	M_{avg}	Alpha+	Alpha−	Gamma+	Gamma−	β_{med}	β_{avg} (r)
Gamma+										
Pearson correlation	.055	.104	.032	−.021	.597**	−.093	1	***.250****	−.068	−.097
Sig. (2-tailed)	.537	.246	.723	.812	.000	.301		***.005***	.450	.287
N	126	126	125	125	126	126	126	126	126	123
Gamma−										
Pearson correlation	−.025	.069	.165	.144	−.133	.326**	.250**	1	−.035	.008
Sig. (2-tailed)	.782	.441	.066	.110	.139	.000	.005		.698	.926
N	126	126	125	125	126	126	126	126	126	123
$\boldsymbol{\beta}_{med}$										
Pearson Correlation	−.065	−.010	.052	.122	−.036	.294**	−.068	−.035	1	.918**
Sig. (2-tailed)	.468	.913	.564	.175	.687	.001	.450	.698		**.000**
N	126	126	125	125	126	126	126	126	126	123
βavg (r)										
Pearson correlation	.006	−.021	.017	.074	−.023	.200*	.001	−.042	.707**	1
Sig. (2-tailed)	.945	.818	.854	.410	.802	.025	.995	.642	.000	
N	126	126	125	125	126	126	126	126	126	123

For better identification by the reader, significant correlations at the 0.05 level are highlighted with an emphasis in bold italics. In addition, for significance at the 0.1 level we added an emphasis in italics. Finally, significant correlations with no statistical interest (such as between average and mean estimations of the same parameter) have been highlighted with an emphasis in bold italics only the significance parameter, not the correlation

** Correlation is significant at the 0.01 level (2-tailed)

* Correlation is significant at the 0.05 level (2-tailed)

Table 8 Correlations among behavioral variables and game indicators

	p_{avg}	p_{vol}	VCC_{ind}	$V_{max,ind}$ (r)	NPL	Q_{avg}	Q_{vol} (r)
E							
Pearson correlation	−.040	−.055	.075	−.109	.075	.002	−.021
Sig. (2-tailed)	.656	.542	.405	.232	.406	.982	.816
N	126	126	126	122	126	126	125
p							
Pearson correlation	−.131	*−.162*	.127	−.069	.002	.096	.071
Sig. (2-tailed)	.143	*.071*	.156	.451	.980	.285	.429
N	126	126	126	122	126	126	125
$\mathbf{M_{med}}$							
Pearson correlation	.105	.057	−.030	.126	−.040	.117	***.198****
Sig. (2-tailed)	.244	.527	.741	.170	.656	.193	***.028***
N	125	125	125	121	125	125	124
$\mathbf{M_{avg}}$							
Pearson Correlation	***.184****	.127	−.110	.120	−.007	.055	.043
Sig. (2-tailed)	***.040***	.159	.224	.191	.935	.545	.634
N	125	125	125	121	125	125	124
Alpha+							
Pearson correlation	−.113	−.030	.021	−.082	.087	***−.242*****	*−.169*
Sig. (2-tailed)	.210	.738	.816	.368	.334	***.006***	*.060*
N	126	126	126	122	126	126	125
Alpha−							
Pearson correlation	.120	*.160*	−.076	.079	−.001	.073	.004
Sig. (2-tailed)	.182	*.073*	.398	.385	.988	.416	.962
N	126	126	126	122	126	126	125
Gamma+							
Pearson correlation	−.122	−.088	.110	*.169*	***.192****	***−.258*****	***−.216****
Sig. (2-tailed)	.173	.325	.221	*.062*	***.031***	***.004***	***.015***
N	126	126	126	122	126	126	125
Gamma−							
Pearson correlation	.061	.077	.119	***.297*****	***.219****	.052	−.060
Sig. (2-tailed)	.500	.394	.186	***.001***	***.014***	.566	.509
N	126	126	126	122	126	126	125
$\boldsymbol{\beta}_{med}$							
Pearson correlation	−.042	−.060	.022	−.055	−.065	−.071	−.085
Sig. (2-tailed)	.640	.502	.804	.547	.470	.427	.349
N	126	126	126	122	126	126	125
$\boldsymbol{\beta}_{avg}$ **(r)**							
Pearson correlation	.021	.003	−.031	−.020	−.056	−.012	.005
Sig. (2-tailed)	.817	.974	.734	.825	.537	.898	.955
N	123	123	123	119	123	123	122

Table 8 continued

	p_{avg}	p_{vol}	VCC_{ind}	$V_{max,ind}$ (r)	NPL	Q_{avg}	Q_{vol} (r)
M_N							
Pearson correlation	.134	.107	−.109	.067	.110	−.020	.048
Sig. (2-tailed)	.135	.237	.228	.463	.220	.823	.594
N	125	125	125	121	125	125	124
M_2							
Pearson correlation	.106	.072	−.052	.105	.016	.071	*.170*
Sig. (2-tailed)	.240	.428	.567	.253	.857	.433	*.060*
N	125	125	125	121	125	125	*124*

Significant correlations at the .05 level are highlighted with an emphasis in bold italic, while for significance at the .1 level we added an emphasis in italic

** Correlation is significant at the .01 level (2-tailed)

* Correlation is significant at the .05 level (2-tailed)

4 Results

After completing the behavioral tests, participants competed in the experimental game. Table 6 summarizes the basic univariate statistics of the indicators that resulted.

Participants exhibited a wide range of strategies: average prices, for instance, range from 12.75 to 18.42 %; $V_{max,ind}$ is close to one, implying that players tended to grant the maximum volume demanded at a given price; and both quality ratios are about 1.0, which means that players were not able to differentiate between low and high quality borrowers (compare this average with the optimal ex post benchmark of about 1.12).

Prior to hypothesis testing, three extreme values were removed for loss aversion (β_{avg} in Table 5). In addition, four observations were removed from two game variables, in accordance with the normality tests and box plots analyzed: three extreme values for $V_{max,ind}$ and one value for the volume-weighted quality ratio (Q_{vol}). These modifications are denoted $\beta_{avg}(r)$, $V_{max,ind}(r)$ and $Q_{vol}(r)$ in what follows. We then conducted three analyses: variable analysis, factorial analysis and cluster analysis (for which results are presented in the subsections that follow). For the sake of simplicity and clarity, we focus only on hypotheses that were satisfied. A description of the tests implemented can be found in the "Appendix".

4.1 Variable analysis

Table 7 shows statistical correlation values for the overconfidence and prospect theory parameters.

For our experimental group, there is evidence that overestimation and overplacement are correlated ($p < .01$). Coming together in both domains are risk seeking (α^+ and α^- are negatively correlated, $p < .05$) and objective weighting of probabilities (γ^+ and γ^- are positively correlated, $p < .01$). There is also strong evidence that loss aversion and risk aversion in the negative domain come together. Finally, in regards to the relationship between overconfidence and prospect theory parameters, we only find positive correlations ($p < 10\ \%$) between α^- and **E**, and between γ^- and **M**. This is a more complex interpretation, as the results suggest that individuals with a more aggressive profile for losses (higher risk seeking and distortion of probabilities) were correlated with lower levels of overconfidence (in terms of overestimation and overprecision). However, this result may also be consistent with Kahneman and Lovallo's (1993) suggestion that biases can cancel out.

Table 8 provides the correlations among behavioral variables and game indicators.

In terms of price and volume of loans (the game decision variables), all significant correlations provide evidence that more aggressive profiles were correlated with riskier credit strategies. In particular, hypotheses 1b and 3a were satisfied to a high degree of significance: γ^- is positively correlated with $V_{\text{max,ind}}$ ($p < .01$), meaning that more risk-seeking participants tended to grant full credit (satisfying hypothesis 1b); and $\mathbf{M}_{\text{avg}}$ is positively correlated with p_{avg} ($p < .05$), suggesting that overprecision in estimating future uncertainty led to a reduced price (satisfying hypothesis 3a). Additionally, for $p < .1$ we obtained three additional results: p is negatively correlated to p_{vol} (satisfying hypothesis 3a); α^- is positively correlated to p_{vol} (a risk seeking profile to avoid sure losses tended to be correlated with more aggressive price policies, satisfying hypothesis 1a); and γ^+ is positively correlated with $V_{\text{max,ind}}$ (satisfying hypothesis 1b).

The more relevant results were obtained for the quality indicators. First, the risk profile for gains has the most powerful ability to predict quality performance. On one hand, γ^+ is correlated positively with the NPL ratio ($p < .05$) and negatively with Q_{avg} ($p < .01$) and Q_{vol} ($p < .05$) (satisfying hypotheses 1c and 1d). On the other hand, α^+ is also negatively correlated with Q_{avg} ($p < .01$) and Q_{vol} ($p < .1$) (satisfying hypothesis 1d). Second, overprecision and risk profile for losses are correlated with quality, since $\mathbf{M}_{\text{med}}$ is positively correlated with Q_{vol} ($p < .05$), suggesting that overprecision reduced quality performance (satisfying hypothesis 3d), and γ^- is positively correlated with the NPL ratio ($p < .05$) (satisfying hypothesis 1c).

These are very suggestive results, as we found twelve pieces of evidence (satisfying six hypotheses) that an aggressive behavioral profile (a high level of overconfidence and more risk seeking behavior) would be significantly correlated with riskier credit strategies, particularly in terms of providing credit to low-quality costumers at a lower price.[11] We found not a single piece of evidence in the opposite direction. It should be mentioned that we could not trace any evidence of loss aversion, a classic in the behavioral literature, explaining credit policies in any direction. Though this is in line with List's (2003) findings that loss aversion does

[11] We checked the robustness of the effects of overprecision comparing the results we obtained under the alternative refinement methods: the estimator M_2 supported that overprecision reduced quality performance (hypothesis 3d) with statistical significance ($p < .1$), while the estimator that assumes normality, M_N, supported that the higher the overprecision the lower the price of credit (hypothesis 3a), but with weak statistical significance ($p = .13$).

Table 9 Regression models: game indicators to behavioral biases

Dependent variable	Model				
	1 p_{avg}	2 $V_{max,ind}$	3 NPL	4 Q_{avg}	5 Q_{vol}
Constant	15.743	.920	5.812	1.091	1.047
E (signific.)	–	–	–	–	–
P (signific.)	–	–	–	–	–
M_{med} (signific.)	–	–	–	–	.075 (.021)
M_{avg}(signific.)	.655 (.043)	–	–	–	–
α^+ (signific.)	–	–	–	–	–
α^- (signific.)	–	–	–	–	–
γ^+ (signific.)	–	–	–	−.108 (.004)	−.082 (.012)
γ^-(signific.)	–	.068 (.001)	.263 (.015)	–	–
β_{med} (signific.)	–	–	–	–	–
β_{avg} (signific.)	–	–	–	–	–
R^2	.034	.088	.048	.066	.089
Adj. R^2	.026	.080	.040	.059	.073

not predict financial behavior in certain contexts, we must be aware that the fact that participants could not lose their own money in this experiment might have made the results insensitive to this variable.

Finally, we conduct a regression analysis, performing a stepwise procedure for variable selection. Results are summarized in Table 9.

The regression results support the main findings of the correlation analysis. First, higher overprecision (lower $\mathbf{M}_{avg}$ or $\mathbf{M}_{med}$) resulted in a more aggressive pricing policy (reduced p_{avg}) and reduced quality (Q_{vol}). Second, higher γ^- (risk seeking behavior for low probability losses) led to a more aggressive volume policy in terms of $V_{max,ind}$ and increased default ratios (NPL ratio). Third, higher γ^+ (risk seeking behavior for medium/high probability gains) reduced quality performance (both Q_{avg} and Q_{vol}).

The game as designed was capable of obtaining evidence on the effects of overprecision and probability weighting on credit policies. In particular, with all relevant information provided in terms of confidence intervals (macroeconomic information and expected default rates) and default probabilities, both correlation and regression analysis showed that excessive precision in estimating future uncertainty and distortion of probabilities bias the credit policies implemented by the players. Moreover, the bias occurred in the expected direction: higher overprecision and a risk-seeking profile led to a more aggressive price–volume policy and reduced quality performance.

Obviously the explanatory power of the models is very low since we are excluding alleged explanatory factors such as expected GDP growth and default rates, but the fact that the coefficient of determination is significantly different from zero highlights the effect that behavioral biases have on credit policies.

Table 10 Factorial analysis: overconfidence

KMO and Bartlett's test	
Kaiser–Meyer–Olkin Measure of Sampling Adequacy	.497
Bartlett's test of sphericity	
Approx. Chi-Sq	82.247
df	3
Sig.	.000
	Component 1
Component matrix[a]	
Overestimation	.919
Overplacement	.907
Overprecision M_{med}	−.208

Extraction method: principal component analysis

[a] 1 components extracted

4.2 Factorial analysis

We test the above hypotheses using correlation and regression analysis but using factors instead of variables. We use principal component analysis (PCA), which aims to build not-directly-observable variables (factors) using directly observable variables (in our research, either the behavioral measures or the game variables). Thus, we firstly make a factorial analysis of the three overconfidence measures, overestimation (**E**), overplacement (**P**) and overprecision (**M**).[12] Results are summarized in Table 10.

The PCA provides intuitive results in that the overconfidence measures are synthetized into one factor, denoted OC, with a coherent interpretation: overestimation and overprecision are positively related to OC, while miscalibration (**M**) is negatively related to OC.

We also conduct a factorial analysis of the five prospect theory parameters considered in our research.[13] Results are summarized in Table 11.

These parameters are assembled into three factors that separate risk profiles for gains, losses, and loss aversion. The rotated component matrix links the first factor to a risk profile for gains (α^+ and γ^+), denoted Gains, with both variables loading positively: thus, the higher the Gains value, the greater the risk seeking.[14] The second factor symmetrically corresponds to the risk profile for losses (α^- and γ^-), denoted Losses. Since both variables load positively, this factor should be interpreted as a higher Losses meaning greater risk aversion.[15] The third factor is loss aversion (β_{med}), combined with some effect by from risk aversion to losses (α^-). Both variables load positively, suggesting that the greater the loss aversion

[12] For simplicity sake, for overprecision in the factorial analysis we only considered the median measure M_{med}.

[13] Again for simplicity sake, for loss aversion in this analysis we only considered the median measure β_{med}.

[14] Note that higher γ^+ implies more risk seeking only for medium/high probabilities.

[15] Again, higher γ^- implies more risk aversion only for medium/high probabilities.

Table 11 Factorial analysis: prospect theory

KMO and Bartlett's test			
Kaiser–Meyer–Olkin Measure of Sampling Adequacy			.424
Bartlett's test of sphericity			
Approx. Chi-Sq			112.082
df			10
Sig.			.000
Rotated component matrix[a]			
	Component		
	1	2	3
Gamma+	.891	.246	−.086
Alpha+	.888	−.229	.013
Gamma−	.082	.922	−.090
Alpha−	−.199	.608	.559
β_{med}	.011	−.065	.933

Extraction method: principal component analysis

Rotation method. Varimax normalization with kaiser

[a] Rotation converged in four iterations

Table 12 Factorial analysis: game indicators

KMO and Bartlett's test		
Kaiser–Meyer–Olkin Measure of Sampling Adequacy		.660
Bartlett's Test of Sphericity		
Approx. Chi-Sq		511.610
df		15
Sig.		.000
Rotated component matrix[a]		
	Component	
	1	2
p_{avg}	.964	.056
p_{vol}	.936	.009
V_{ind}	−.904	−.065
NPL	.078	−.739
Q_{avg}	.071	.900
Q_{vol} (r)	.158	.907

Extraction method: principal component analysis

Rotation. Varimax Normalization with Kaiser[a]

[a] Rotation converged in three iterations

(and risk aversion in the negative domain), the higher the factor value. Hence, we denote this factor Loss_aversion.

We now conduct a factorial analysis of all seven game indicators. In a first stage we obtain a factorization that is incoherent in terms of $V_{max,ind}$ interpretation. Excluding $V_{max,ind}$ from the factorial analysis, in a second stage we obtain the results summarized in Table 12.

The PCA posits two factors, with the Kaiser–Meyer–Olkin (KMO) measure improving significantly (from .489 to .660) and with both factors synthetizing

Table 13 Correlations among factors

	OC	Gains	Losses	Loss_av
Strategy				
Pearson correlation	−.097	−.096	.071	.018
Sig. (2-tailed)	.284	.288	.428	.845
N	124	125	125	125
Quality				
Pearson correlation	−.012	***−.266****	−.069	−.040
Sig. (2-tailed)	.897	***.003***	.445	.657
N	124	125	125	125

For better identification by the reader, significant correlations at the .05 level are highlighted with an emphasis in bold italics

information in a coherent manner. The first factor synthetizes price and volume, which were the variables that determined a participant's strategy. From the way they load—with price indicators positively related and V moving in the opposite direction—we interpret that the higher the factor value, the more conservative (higher price, lower volume) the participant's strategy. We call this factor Strategy, interpreted as the higher the factor, the more conservative the strategy. The second factor synthetizes all indicators that determine the quality profile of the credit policy, hence we denote it Quality. Our intuition was correct about how the indicators load: the ratio of non-performing loans to total loans is negatively related to Quality while Q_{avg} and Q_{vol} are positively related to Quality.

Taking these factors into consideration to further test the above hypotheses, we again conduct correlation and regression analyses of credit factors against behavioral factors. The correlation analysis is summarized in Table 13.

Correlations provide additional evidence that the risk profile for gains affects the quality performance of credit policies. In particular, consistent with our hypothesis, participants with greater risk seeking for gains tended to grant lower prices to lower quality clients. Regression analysis yields similar results, but attributing a causality effect to γ^+ in particular (Table 14).

4.3 Cluster analysis

The statistical analysis of the experimental results is completed with a cluster analysis. Different clustering alternatives were analyzed, considering different methods and analyses for both variables and factors. Easiest to interpret was clustering in terms of game factors,[16] Strategy and Quality, with six different clusters obtained. Results are shown in Fig. 4.

We may characterize the irrationality of these subgroups (somehow in the manner of Forsythe et al. 1992) in terms of their Strategy and Quality values. Thus, the largest cluster (cluster 1) centered both in terms of Strategy and Quality, while the smaller groups reflect sparser credit policies. Cluster 2 would tend to exhibit the

[16] We used 125 observations as we excluded one outlier from variable Q_{vol} (which loads on Quality). Additionally, one more observation is lost for OC, since we had a missing value for M ratios from the beginning.

Table 14 Regression models: game factors to behavioral factors and variables

REGRESSION MODEL: GAME FACTORS TO BEHAVIORAL FACTORS

Dependent variable (Game factor)	Model 1 Strategy	Model 2 Quality
Constant	-	-.016
OC (signif.)	-	-
Gains (signif.)	-	-.261 (.003)
Losses (signif.)	-	-
Loss Aversion (signif.)	-	-
R^2	-	.071
adj. R^2	-	.063

REGRESSION MODEL: GAME FACTORS TO BEHAVIORAL VARIABLES

Dependent variable (Game factor)	Model 1 Strategy	Model 2 Quality
Constant	-	.645
E (signif.)	-	-
P (signif.)	-	-
M_{med} (signif.)	-	-
M_{avg} (signif.)	-	-
α^+ (signif.)	-	-
α^- (signif.)	-	-
γ^+ (signif.)	-	-1.029 (.003)
γ^- (signif.)	-	-
β_{med} (signif.)	-	-
β_{avg} (signif.)	-	-
R^2	-	.072
adj. R^2	-	.064

most conservative Strategy (i.e., higher prices and lower credit volumes) and an average Quality, while cluster 4 would be in the opposite place (an average Quality for a riskier Strategy). Most participants in cluster 5 would tend to follow a Strategy which is only slightly aggressive, but when it comes to screening high-quality clients from low-quality ones, they exhibit a low Quality (higher NPL ratios and a tendency to grant lower prices to riskier clients). Cluster 3 would be in the opposite place in terms of Quality (i.e., a tendency to be excessively conservative with lower quality clients) for a neutral Strategy overall. Finally, cluster 6 is formed of a single participant who followed an aggressive Strategy (lower prices overall) but with a higher Quality, which implies excessive conservatism with low-quality clients. This suggests this player must have granted too low prices mostly to high-quality clients.

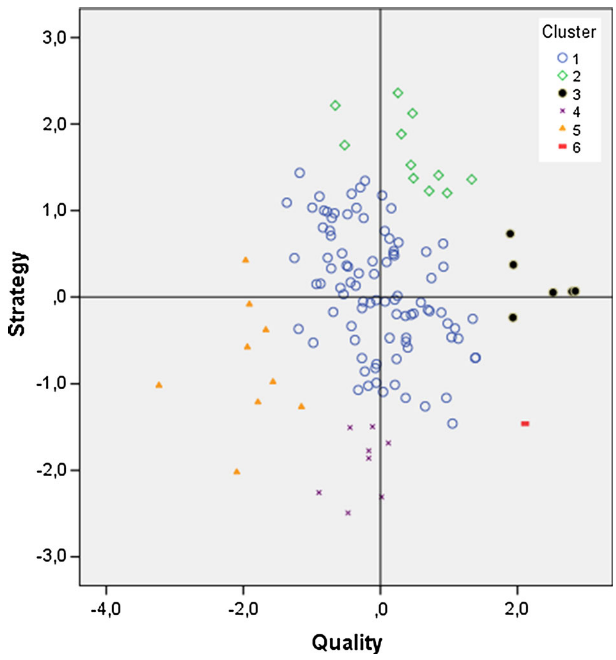


Fig. 4 Cluster analysis

Table 15 summarizes the descriptive statistics (means and standard error of the means) for each cluster in terms of the six behavioral and game factors.

However the deviation of some subgroups in terms of credit quality, price and volume is easy to characterize, it is difficult to go further to interpret which are the behavioral biases behind these clusters. We may use dispersion diagrams (Fig. 5) based on the mean values show how the different behavioral profiles of each cluster might affect credit strategies. We emphasize careful interpretation, however, as any statistical analysis based on correlations or causality effects are unsound at this level.[17]

The largest cluster (cluster 1) would not only include players that exhibited centered Strategy and Quality factors, but would also indicate that these were generally neutral in terms of any of the four behavioral factors summarizing overconfidence and risk profile. Consequently, it is the behavior of the smaller groups which needs to be explained: what would happen if a group is biased? Would their Strategies and Quality vary in a predictable manner? We conduct an ANOVA test for differences of means across clusters to obtain significant evidence that the clusters are different in terms of Gains ($p < .05$).

This may be interpreted in two instances. First, since the distortion of probabilities in the positive domain, γ^+, loads on this factor, the effect of this

[17] Having only 6 observations (clusters) would itself invalidate the statistical significance of any correlations or regression analyses. Moreover, much information is lost when we use average values to be representative of all individual observations in a cluster.

Table 15 Cluster: descriptive statistics

Cluster	N	OC		Gains		Losses		Lossavers		Strategy		Quality	
		Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error	Mean	Std. Error
1	90[a]	.008	.108	−.114	.099	.024	.105	−.079	.105	.073	.074	−.035	.069
2	11	−.302	.196	−.218	.323	−.155	.282	.398	.330	1.676	.125	.422	.178
3	6	−.215	.411	.104	.396	.101	.468	-.415	.210	.175	.136	2.320	.184
4	8	−.008	.491	.945	.491	−.453	.415	.172	.313	−1.922	.135	−.267	.115
5	9	.324	.277	.515	.269	.252	.337	.400	.402	−.790	.241	−1.925	.188
6	1	−.184		−.342		−.224		−.657		−1.461		2.110	

[a] Data for OC includes 89 observations

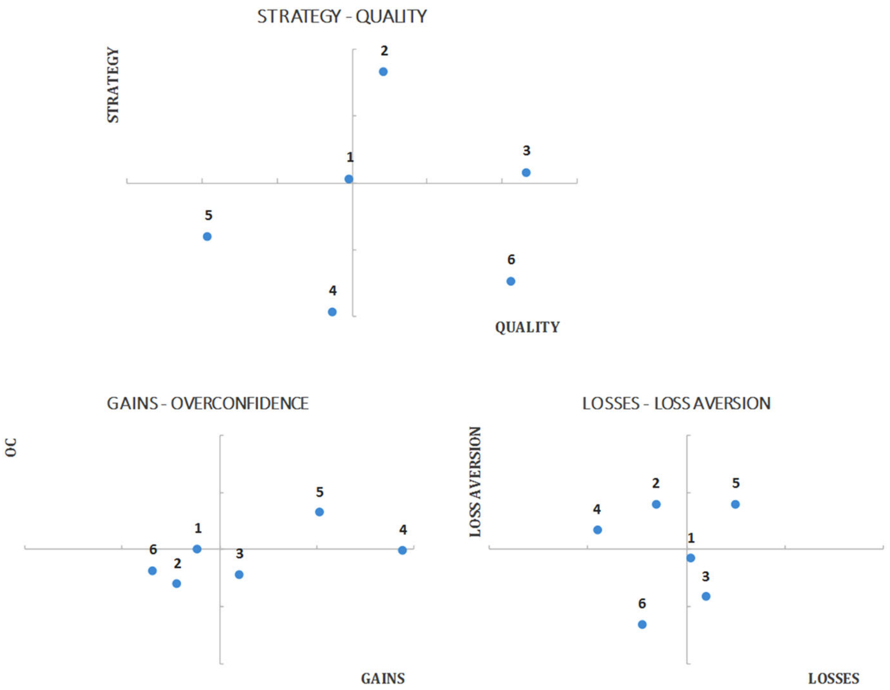


Fig. 5 Cluster analysis. Dispersion diagrams for average values

feature of prospect theory would be the most solid of our research, since the four statistical methods implemented (correlations, regressions, factorial analysis and cluster analysis) provide a positive evidence in that sense. Second, the asymmetry in incentives in our experiment, where participants may win a prize but never lose their own money, has probably make it easier to obtain evidence of the effects of risk seeking in the positive domain (Gains). Thus, the absence of any evidence in the factorial and cluster analyses of the effects on the negative domain (Losses) as well as in terms of loss aversion does not necessarily mean they do not exist, but that our experiment is limited to obtain evidence of it.

5 Concluding remarks

We designed an original business simulation game that replicates basic decision-making processes for a bank granting credit to clients under conditions of uncertainty and risk. In order to test whether overconfidence and prospect theory are able to explain excessive lending patterns by retail banks, we organized a series of experimental sessions with 126 under- and postgraduate students that was divided in two parts. The first part consisted of a series of short tests to measure participants' level of overconfidence and risk profile according to prospect theory. The second part was the simulation game itself. We tested several hypotheses about the effects of risk seeking, loss aversion and overconfidence, with main results as follows.

First, we found extensive evidence that an aggressive behavioral profile—defined as high levels of overconfidence and risk seeking—is correlated with riskier credit strategies, particularly in terms of providing credit to low-quality clients at a lower price. We found not a single piece of evidence in the contrary direction, neither were we are able to locate evidence that loss aversion has any effects on credit policies.

Second, overestimation and overplacement were only weakly observed in our experiment, despite the fact that their effects on financial decision making have been documented in the literature (e.g., Graham et al. 2009; Grinblatt and Keloharju 2009; Deaves et al. 2009). Instead, our game design revealed helpful to obtain evidence of overprecision and probability weighting effects on credit policies: all players were given identical information to play the game, expressed in terms of confidence intervals and default probabilities, and it is in terms of overprecision (related to confidence intervals) and probability weighting where the most solid results were obtained. In particular, our results suggest that overprecision in estimating future uncertainty and the risk profile for gains (mostly attributable to a distortion of probabilities) foster lower prices and higher volumes of credit granted, and reduce quality. These results support the classic finding that miscalibration has an impact on financial decision making, as in theoretical models such as that of Odean (1998), which suggests that investors trade too much, and the experimental research by Biais et al. (2005) and Deaves et al. (2009), who observed that miscalibration reduces trading profits. This effect is complemented by the evidence that a distortion of probabilities under prospect theory captures the common preference for a lottery-like wealth distribution (Barberis and Huang 2008), which explains empirical facts like people attributing too much weight to rare events, and stocks that are expected to be positively skewed having lower expected returns (Boyer et al. 2010). Indeed, the most consistent result in our experiment was that of distortion of probabilities fostered lower loan prices to potential borrowers of a lower quality.

Third, the effects of these biases on the quality of credit tend to favor the external validity of our results. Participants in the experiment were aware that their strategies were to be measured and scrutinized in terms of credit prices and volumes. Hence, the indicators might well be affected by their strategic behavior: e.g., participants behind in the game (losing in terms of profits) might make weird decisions because they had nothing to lose. However, participants were not aware that their behavior in terms of quality was also being scrutinized. This is good news, since the most significant results of the experiment were regarding the effects of overprecision and probability distortion over quality performance.

This paper represents a first effort to explore the potential effects of behavioral biases on credit policies. Time constraints to implement both a series of psychological tests and the simulation game imposed the restriction that only one repetition of the game was possible.[18] If several rounds of the game with randomized economic scenarios were implemented, the external validity of the

[18] Participants in the experimental sessions spent an average of three hours completing the overconfidence and prospect theory tests and the simulation game, instructions included.

results would undoubtedly be enhanced. This would help mitigate the effects of strategic behavior and avoid any randomization bias (Viera and Bangdiwala 2007); it would also admit the possibility of testing the debiasing effects of learning and experience —which some studies suggest tend to mitigate observed deviations from rationality (e.g., List 2003; van de Kuilen and Wakker 2006).

Finally, some limitations of this experiment need to be resolved in future replications of the experiment. Firstly, it would be interesting to extend the way we tested the behavioral biases of the participants to other tests and elicitation methods available, such as cumulative prospect theory, nonparametric methods, alternative measures of overconfidence (see Hilton et al. 2011, for a recent review) and alternative methods to elicit the measures used (see, for instance, Glaser et al. 2013). Secondly, the robustness of our results supporting the effects of overprecision on credit policies should be further qualified. In future research we plan to enhance the elicitation method to obtain measures that are more stable at the individual level for different refinement methods. Lastly, the monetary incentives in the game and the absence of penalties for excessive risk taking may have biased the results. This is, unfortunately, a common drawback in this kind of experimental research, since few people are willing to participate in experiments in which they stand to lose real money, although in future versions of this game we should try to overcome this problem. Nonetheless, we defend the results obtained in two ways: first, the most conclusive results were in terms of quality, a variable that participants were not aware of; and second, the combination of incentives and moral hazard when costs are not borne resembles the alleged behavior of bank executives during the recent financial crisis.

Acknowledgments The authors wish to thank the Editor of Review of Managerial Science and two anonymous referees for their insightful comments and suggestions. They also thank Andrea Ceschi and Paulino Martínez for very valuable support in the experiment design, and Enrico Cervellati, Xosé Manuel M. Filgueira and Rafael García for technical assistance. M.A. appreciates the financial aid received from the Galician Autonomous Government through project GPC 2013-045.

Appendix

IBM SPSS Statistics version 21 was used for the statistical analysis. Technical specifications for the analyses and raw data for the strategies implemented by participants in the experiment are described in the Supplementary Material to this paper.

References

Abdellaoui M, Bleichrodt H, L'Haridon O (2008) A tractable method to measure utility and loss aversion under prospect theory. J Risk Uncertain 36:245–266

Abdellaoui M, Bleichrodt H, Kammoun H (2013) Do financial professionals behave according to prospect theory? An experimental study. Theory Decis 74(3):411–429

Acharya V, Naqvi H (2012) The seeds of a crisis: a theory of bank liquidity and risk taking over the business cycle. J Financ Econ 106:349–366
Altman EI (1968) Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J Finance 23(4):589–609
Argenti J (1976) Corporate collapse: the causes and symptoms. McGraw-Hill, New York
Asch SE (1952) Social psychology. Prentice Hall, New York
Barber BM, Odean T (2001) Boys will be boys: gender, overconfidence, and common stock investment. Q J Econ 116(1):261–292
Barber BM, Odean T (2002) Online investors: do the slow die first? Review of Financial Studies 15:455–489
Barberis N, Huang M (2008) Stocks as lotteries: the implications of probability weighting for security prices. Am Econ Rev 98(5):2066–2100
Beaver W (1966) Market prices, financial ratios and the prediction failure. Journal of Accounting Research 6:179–192
Benartzi S, Thaler RH (1995) Myopic loss aversion and the equity premium puzzle. Q J Econ 110(1):73–92
Berger AN, Mester LJ (1997) Inside the black box: what explains differences in the efficiencies of financial institutions? J Bank Finance 21:895–947
Berger AN, Kick T, Schaeck K (2012), Executive board composition and bank risk taking, Deutsche Bundesbank Discussion Paper No. 03/2012
Biais B, Hilton D, Mazurier K, Pouget S (2005) Judgmental overconfidence, self-monitoring and trading performance in an experimental financial market, Rev Econ Stud 287–312
Boyer BH, Mitton T, Vorkink K (2010) Expected idiosyncratic skewness. Rev Financ Stud 23(1):169–202
Burnside C, Han B, Hirsleifer D, Wang TY (2011) Investor overconfidence and the forward premium puzzle. Rev Econ Stud 78:523–558
Camerer C, Lovallo D (1999) Overconfidence and excess entry: an experimental approach. Am Econ Rev 89(1):306–318
Daniel KD, Hirshleifer D, Subrahmanyam A (1998) Investor psychology and security market under- and overreactions. J Finance 53(6):1839–1885
De Freitas S, Oliver M (2006) How can exploratory learning with games and simulations within the curriculum be most effectively evaluated? Comput Educ 46(3):249–264
Deaves R, Luders E, Luo GY (2009) An experimental test of the impact of overconfidence and gender on trading activity. Rev Finance 13(3):555–575
Demsetz H (1973) Industry structure, market rivalry and public policy. J Law Econ 16(1):1–9
Deshmukh S, Goel AM, Howe KM (2010) CEO overconfidence and dividend policy. SSRN eLibrary
Djankov S, McLiesh C, Shleifer A (2005) Private credit in 129 countries, NBER Working Paper Series, w11078. Available at SSRN: http://ssrn.com/abstract=652366
Fahlenbrach R, Stulz RM (2011) Bank CEO incentives and the credit crisis. J Financ Econ 99(1):11–26
Forsythe R, Nelson F, Neumann GR, Wright J (1992) Anatomy of an experimental political stock market. Am Econ Rev 82(5):1142–1161
Glaser M, Weber M (2007) Overconfidence and trading volume. Geneva Risk Insur Rev 32:1–36
Glaser M, Langer T, Weber M (2013) True overconfidence in interval estimates: evidence based on a new measure of miscalibration. J Behav Decis Mak 26:405–417
Graham JR, Harvey CR, Huang H (2009) Investor competence, trading frequency, and home bias. Manage Sci 55(7):1094–1106
Gredler ME (2004) Games and simulations and their relationships to learning. In: Jonassen DH (ed) Handbook of research on educational communications and technology, vol 2. Erlbaum, Mahwah, pp 571–581
Grinblatt M, Keloharju M (2009) Sensation seeking, overconfidence and trading activity. J Finance 64(2):549–578
Hens T, Bachmann K (2008) Behavioural finance for private banking. Wiley, New York
Hilton D, Régner I, Cabantous L, Charalambides L, Vautier S (2011) Do positive illusions predict overconfidence in judgment? A test using interval production and probability evaluation measures of miscalibration. J Behav Decis Mak 24:117–139
Huang S-Y, Lambertides N, Steeley JM (2012) Cash hoards and managerial overconfidence. In: 19th Annual Conference of the Multinational Finance Society, Krakow, June 2012

Kahneman D, Lovallo D (1993) Timid choices and bold forecasts: a cognitive perspective on risk taking. Manage Sci 39(1):17–31
Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47(2):263–291
Keen S (2011) A monetary Minsky model of the great moderation and the great recession. J Econ Behav Organ. doi:10.1016/j.jebo.2011.01.010
Keys BJ, Mukherjee T, Seru A, Vig V (2010) Did securitization lead to lax screening? Evidence from subprime loans. Q J Econ 125(1):307–362
Knetsch JL (1989) The endowment effect and evidence of nonreversible indifference curves. Am Econ Rev 79(5):1277–1284
Levitt SD, List JA (2007) What do laboratory experiments measuring social preferences tell us about the real world? J Econ Perspect 21(2):153–174
Levitt SD, List JA (2009) Field experiments in economics: the past, the present, and the future. Eur Econ Rev 53:1–18
List JA (2003) Does market experience eliminate market anomalies? Q J Econ 118:41–71
List JA (2009) An introduction to field experiments in economics. J Econ Behav Organ 70:439–442
Lunenburg FC (2010) Group decision making: the potential for groupthink. Int J Manag Bus Adm 13(1):1–6
Malmendier U, Tate G (2005a) CEO overconfidence and corporate investment. J Finance 60(6):2661–2700
Malmendier U, Tate G (2005b) Does overconfidence affect corporate investment? CEO Overconfidence measures revisited. Eur Financ Manag 11(5):649–659
Milgram S (1963) Behavioral study of obedience. J Abnorm Soc Psychol 67(4):371–378
Milgram S (1974) Obedience to authority: an experimental view. Harpercollins, New York
Moore PG (1977) The manager's struggle with uncertainty. J R Stat Soc Series A 149:129–165
Moore DA, Healy PJ (2008) The trouble with overconfidence. Psychol Rev 115(2):502–517
Odean T (1998) Volume, volatility, price and profit when all traders are above average. J Finance 53(6):1887–1934
Odean T (1999) Do investors trade too much? Am Econ Rev 89(5):1279–1298
Peón D, Calvo A (2012) Using behavioral economics to analyze credit policies in the banking industry. Eur Res Stud 15(3):135–144
Peón D, Calvo A, Antelo M (2014) A short test for overconfidence and prospect theory: an experimental validation, MPRA working paper no. 57899. Available at http://mpra.ub.uni-muenchen.de/57899/
Peón D, Antelo M, Calvo A (2015) On informational efficiency of the banking sector: a behavioral model of the credit boom. Stud Econ Finance 32(2), forthcoming
Prelec D (1998) The probability weighting function. Econometrica 66(3):497–527
Presbitero AF, Udell GF, Zazzaro A (2012) The home bias and the credit crunch: a regional perspective, MoFiR workshop on banking, VoxEu.org 12-feb-2012. http://www.voxeu.org/index.php?q=node/7614
Rabin M (2000) Risk aversion and expected-utility theory: a calibration theorem. Econometrica 68(5):1281–1292
Rieger MO, Wang M (2008) Prospect theory for continuous distributions. J Risk Uncertain 36:83–102
Rodríguez M (2000) Métodos y modelos de pronóstico del fracaso empresarial: Una aproximación empírica a la realidad empresarial de la CC.AA de Galicia, Doctoral thesis, Universidade da Coruña
Roll R (1986) The hubris hypothesis of corporate takeovers. J Bus 59:197–216
Rötheli TF (2012) Oligopolistic banks, bounded rationality, and the credit cycle. Econ Res Int. doi:10.1155/2012/961316
Samuelson W, Zeckhauser R (1988) Status quo bias in decision making. J Risk Uncertain 1(1):7–59
Scheinkman JA, Xiong W (2003) Overconfidence and speculative bubbles. J Polit Econ 111(6):1183–1219
Shefrin H (2008) Ending the management illusion: How to drive business results using the principles of behavioral finance, 1st edn. Mc-Graw Hill, New York
Shefrin H, Cervellati E (2011) BP's failure to debias: underscoring the importance of Behavioral Corporate Finance, WP available at SSRN. http://ssrn.com/abstract=1769213
Shefrin H, Statman M (1985) The disposition to sell winners too early and ride losers too long: theory and evidence. J Finance 40:777–790
Shiller RJ (1984) Stock prices and social dynamics. Brookings papers on economic activity, Fall, pp 457–498

Smith VL (2001) Experimental economics. In: Smelser NJ, Baltes PB (eds) International Encyclopedia of the Social & Behavioral Sciences. Elsevier, Amsterdam, pp 5100–5108

Soll JB, Klayman J (2004) Overconfidence in interval estimates. J Exp Psychol Learn Mem Cogn 30(2):299–314

Stiglitz JE, Weiss A (1981) Credit rationing in markets with imperfect information. Am Econ Rev 71:393–410

Thaler RH (1980) Toward a positive theory of consumer choice. J Econ Behav Organ 1(1):39–60

Thaler RH, Johnson EJ (1990) Gambling with the house money and trying to break even: the effects of prior outcomes on risky choice. Manag Sci 36(6):643–660

Townsend R (1979) Optimal contracts and competitive markets with costly state verification. J Econ Theory 21:265–293

Tversky A, Kahneman D (1991) Loss aversion in riskless choice: a reference-dependent model. Q J Econ 106(4):1039–1061

Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain 5(4):297–323

van de Kuilen G, Wakker PP (2006) Learning in the Allais paradox. J Risk Uncertain 33:155–164

Viera AJ, Bangdiwala SI (2007) Eliminating bias in randomized controlled trials: importance of allocation concealment and masking. Fam Med 39(2):132–137