Identification of Preferences and Expectations in a Natural Experiment

by
Fabrizio Botti, Anna Conte, Daniela T. Di Cagno, Carlo D'Ippoliti and Peter G. Moffatt

LUISS Guido Carli; e-mail: fbotti@luiss.it
LUISS Guido Carli, University of Rome I "La Sapienza" and University of Rome II "Tor Vergata"; e-mail: aconte@luiss.it
LUISS Guido Carli; e-mail: ddicagno@luiss.it
LUISS Guido Carli and University of Rome "La Sapienza": cdippoliti@luiss.it
University of East Anglia: P.Moffatt@uea.ac.uk

AFFARI TUOI

The game starts with 20 players, one from each of the 20 Italian regions.

They are each randomly assigned a sealed box,

containing one of the following prizes:

Prizes as displayed to players

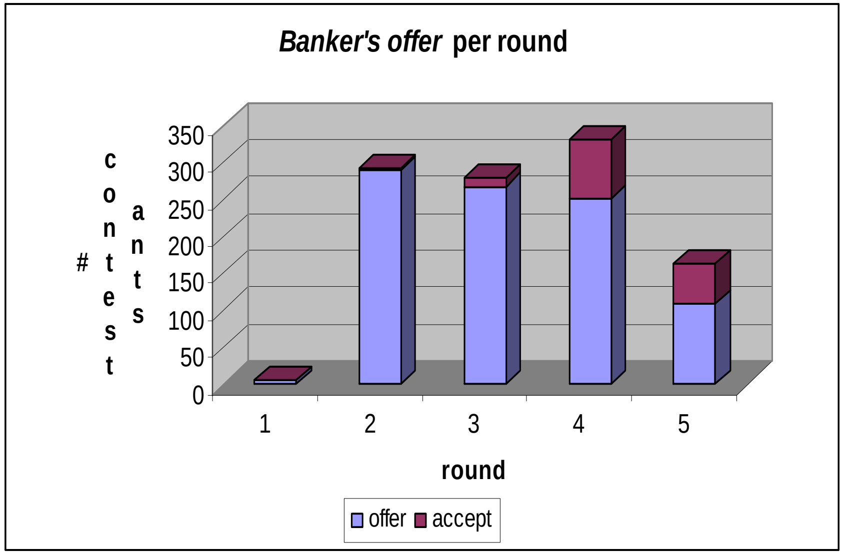
€ 0.01	€ 5,000
€ 0.20	€ 10,000
€ 0.50	€ 15,000
€ 1	€ 20,000
€ 5	€ 25,000
€ 10	€ 50,000
€ 50	€ 75,000
€ 100	€ 100,000
€ 250	€ 250,000
€ 500	€ 500,000

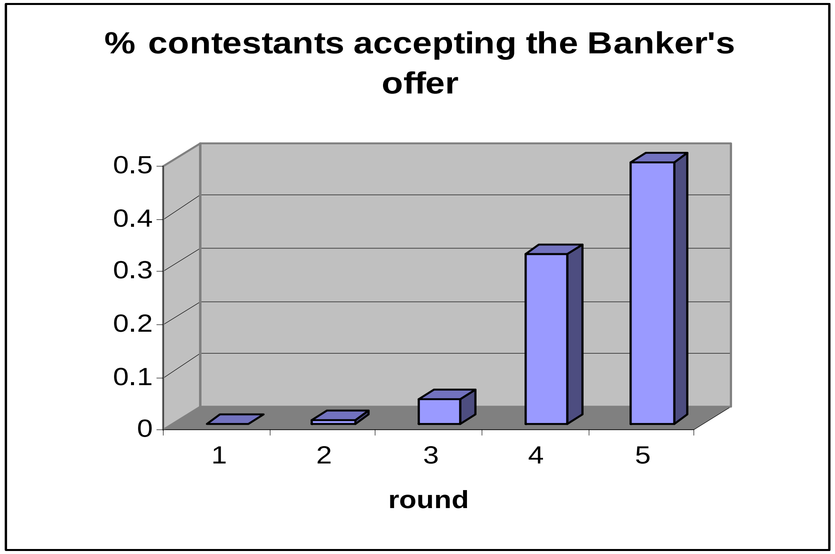
The show begins by contestants answering a general knowledge question. The first contestant to answer correctly is selected to play against the *Banker*.

In each of the 5 rounds, the contestant opens a fixed number of boxes (6 in the first round, then groups of 3 boxes); on each occasion when a box is opened the cash value of that box is revealed, indicating a sum of money which is no longer available to the contestant.

Between every two rounds, the *Banker* makes an offer: he either offers the contestant the opportunity to <u>change her box</u> with any of the remaining boxes, or offers a <u>certain amount of money</u> to quit the game. If the player accepts the money, the game ends; otherwise she proceeds to the next round.

If the contestant gets to the final round without having accepted any of the *Banker's offers*, she wins the content of her selected box.





THE SAMPLE

Our sample consists of 294 showings, and therefore contains data on 294 contestants' decisions.

We limit our analysis to the last 3 rounds for various reasons:

- offers are rarely accepted in rounds 1 and 2; this implies that there is not enough variability in the data to explain the choice process;
- we notice that contestants are not taking the Game seriously in the first two rounds; in these stages, anything can happen, and they prefer to stay in the game whatever the offer is.
- the audience tend to participate actively in the first two rounds, offering advice to the contestant. In later stages, the audience appears to respect the contestant's choices more.

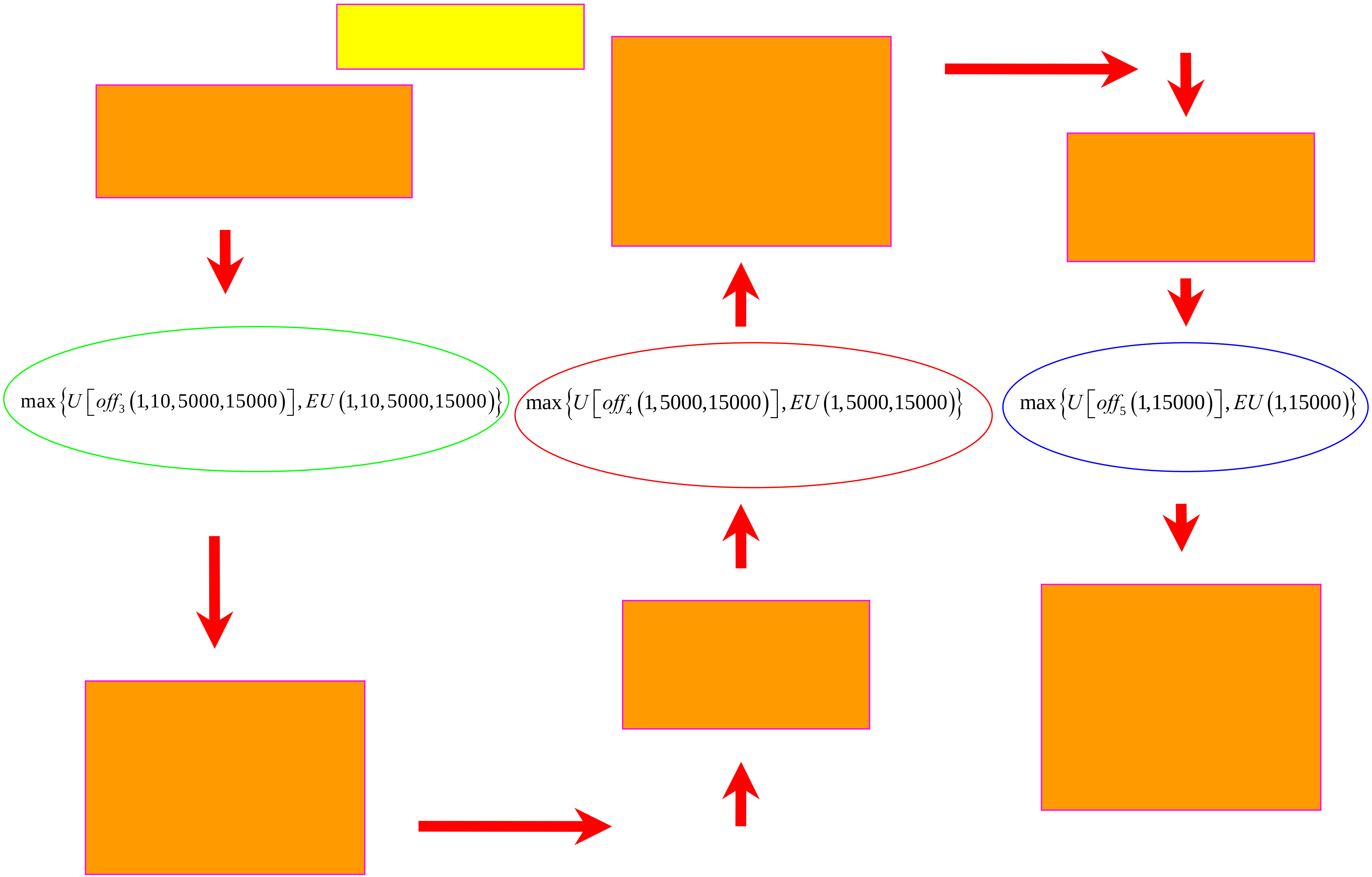
MODELLING CHOICES: MYOPIC AND FORWARD-LOOKING CONTESTANTS

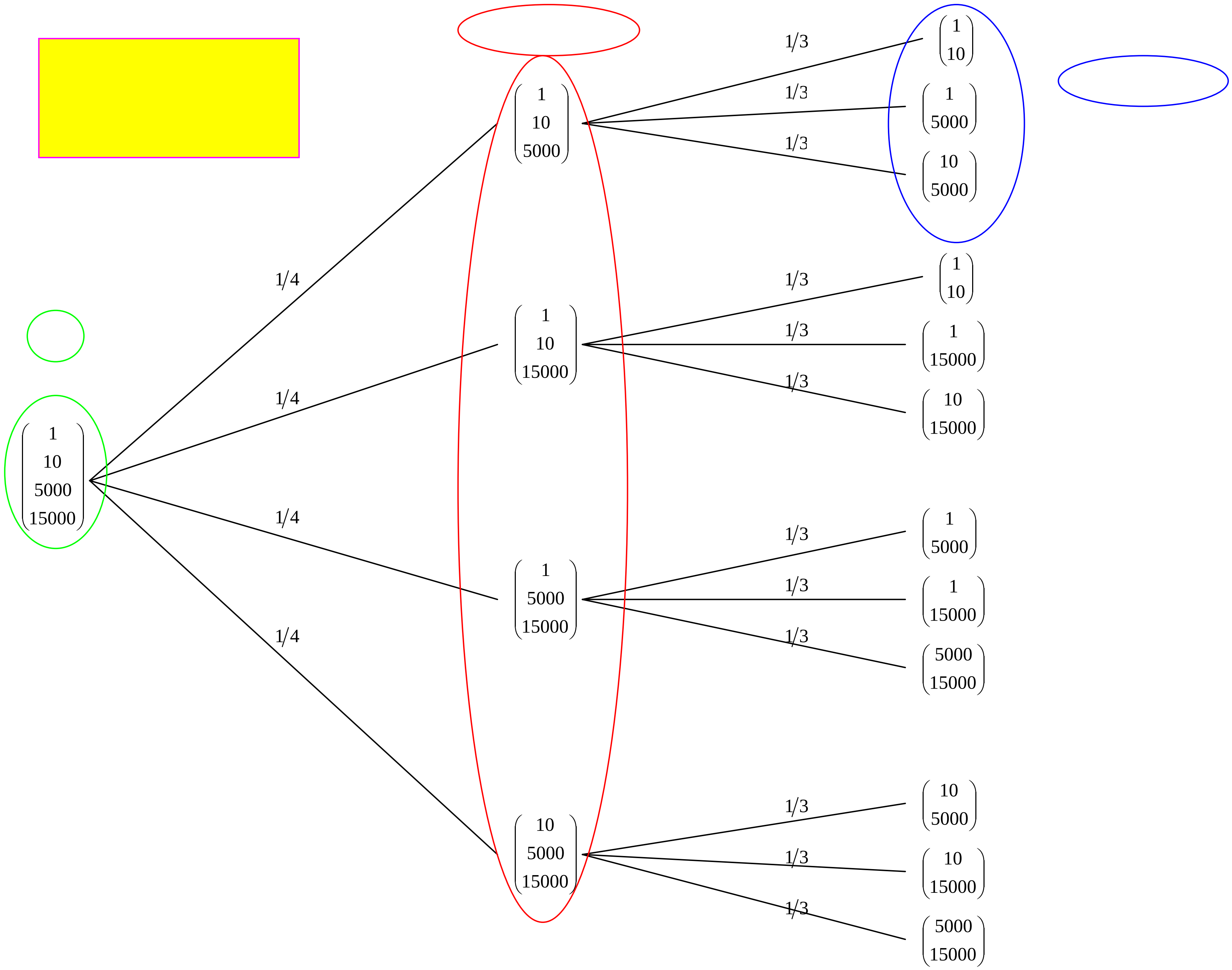
In the ultimate stage of the game, the contestant is offered the choice between participating in a lottery with two equally probable outcomes (the remaining two prizes), or to accept the final offer made by the *Banker*.

In the first four rounds, we need to make a distinction between myopic contestants and forward-looking contestants:

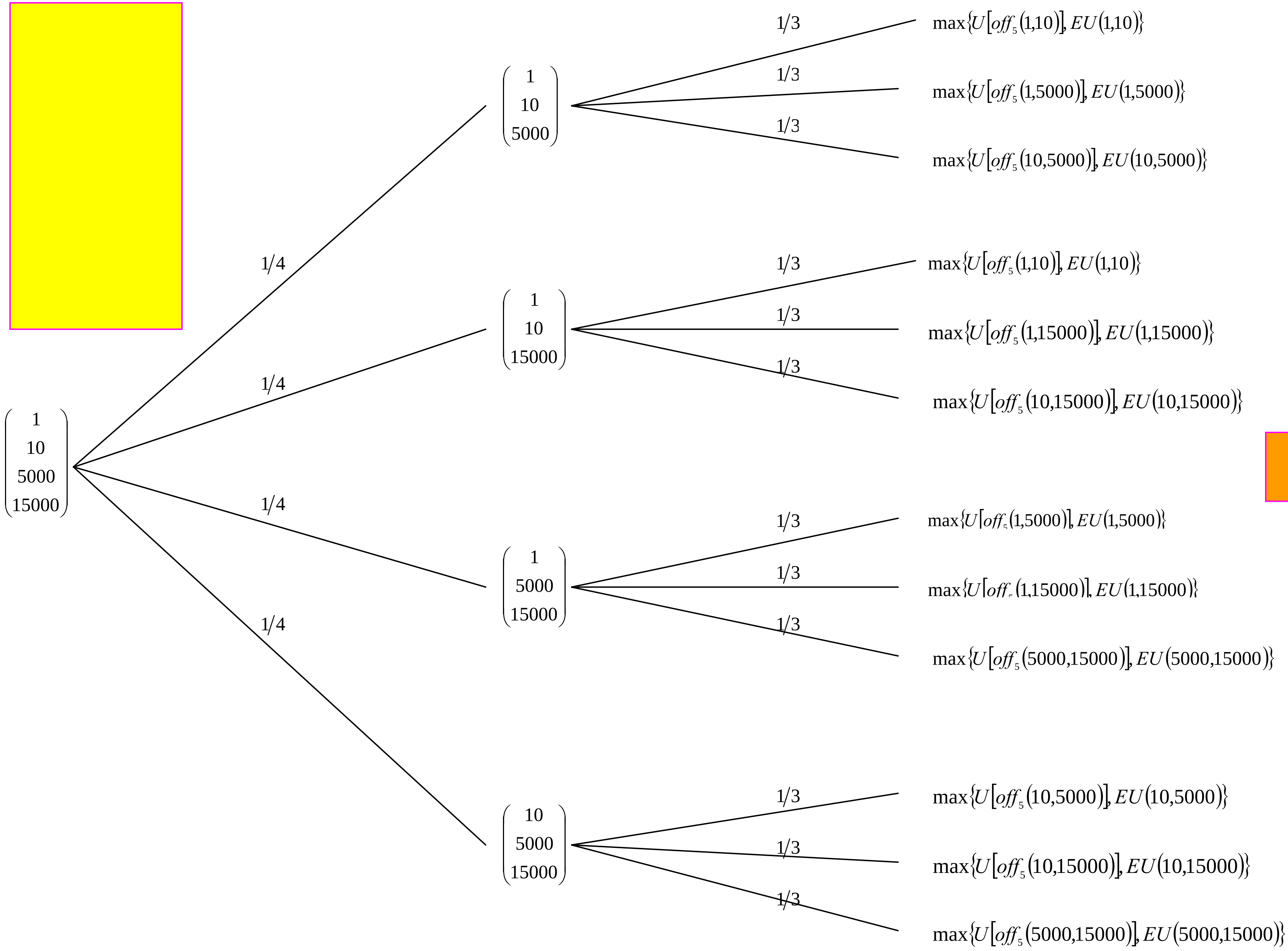
- if the contestant behaves myopically, she will choose between the lottery and *Banker*'s offer without taking into account the subsequent rounds, and in particular the prospect of getting a higher offer later on.

- forward-looking contestants consider all the possible lotteries they might be confronted with in future rounds, because they realize that they are not playing a one-shot lottery but a sequence of nested lotteries and that they might get a higher offer in later rounds. In making decisions, they need to form expectations of the *Banker's Offer* that will arise under every possible lottery that might be faced in future rounds.


$\max\left\{U\left[off_3\left(1,10,5000,15000\right)\right], EU\left(1,10,5000,15000\right)\right\}$
$\max\left\{U\left[off_4\left(1,5000,15000\right)\right], EU\left(1,5000,15000\right)\right\}$
$\max\left\{U\left[off_5\left(1,15000\right)\right], EU\left(1,15000\right)\right\}$



1
10
5000
15000
1/4
1/4
1/4
1/4
1
10
5000
1
10
15000
1
5000
15000
10
5000
15000
1/3
1/3
1/3
1
10
1
5000
10
5000
1
15000
10
15000
5000
15000



(1, 10, 5000, 15000)
1/4
1/4
1/4
1/4
(1, 10, 5000)
(1, 10, 15000)
(1, 5000, 15000)
(10, 5000, 15000)
1/3
$\max\{U[off_5(1,10)], EU(1,10)\}$
$\max\{U[off_5(1,5000)], EU(1,5000)\}$
$\max\{U[off_5(10,5000)], EU(10,5000)\}$
$\max\{U[off_5(1,10)], EU(1,10)\}$
$\max\{U[off_5(1,15000)], EU(1,15000)\}$
$\max\{U[off_5(10,15000)], EU(10,15000)\}$
$\max\{U[off_5(1,5000)], EU(1,5000)\}$
$\max\{U[off_5(1,15000)], EU(1,15000)\}$
$\max\{U[off_5(5000,15000)], EU(5000,15000)\}$
$\max\{U[off_5(10,5000)], EU(10,5000)\}$
$\max\{U[off_5(10,15000)], EU(10,15000)\}$
$\max\{U[off_5(5000,15000)], EU(5000,15000)\}$

$$\begin{pmatrix}1\\10\\5000\\15000\end{pmatrix}$$

$$\xrightarrow{1/4}\ \max\left(U\left[off_4(1,10,5000)\right]\ ,\ \begin{pmatrix} \frac{1}{3}\max\left\{U\left[off_5(1,10)\right], EU(1,10)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(1,5000)\right], EU(1,5000)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(10,5000)\right], EU(10,5000)\right\} \end{pmatrix}\right)$$

$$\xrightarrow{1/4}\ \max\left(U\left[off_4(1,10,15000)\right]\ ,\ \begin{pmatrix} \frac{1}{3}\max\left\{U\left[off_5(1,10)\right], EU(1,10)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(1,15000)\right], EU(1,15000)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(10,15000)\right], EU(10,15000)\right\} \end{pmatrix}\right)$$

$$\xrightarrow{1/4}\ \max\left(U\left[off_4(1,5000,15000)\right]\ ,\ \begin{pmatrix} \frac{1}{3}\max\left\{U\left[off_5(1,5000)\right], EU(1,5000)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(1,15000)\right], EU(1,15000)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(5000,15000)\right], EU(5000,15000)\right\} \end{pmatrix}\right)$$

$$\xrightarrow{1/4}\ \max\left(U\left[off_4(10,5000,15000)\right]\ ,\ \begin{pmatrix} \frac{1}{3}\max\left\{U\left[off_5(10,5000)\right], EU(10,5000)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(10,15000)\right], EU(10,15000)\right\}+ \\ \frac{1}{3}\max\left\{U\left[off_5(5000,15000)\right], EU(5000,15000)\right\} \end{pmatrix}\right)$$

$$
EU^{forward}\left(X_{i3}\right)=\max\left(U\left[off_3\left(1,10,5000,15000\right)\right],\left(\begin{array}{l}
\frac{1}{4}\max\left(U\left[off_4\left(1,10,5000\right)\right],\left(\begin{array}{l}
\frac{1}{3}\max\left\{U\left[off_5\left(1,10\right)\right],EU\left(1,10\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(1,5000\right)\right],EU\left(1,5000\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(10,5000\right)\right],EU\left(10,5000\right)\right\}
\end{array}\right)\right)+\\
\frac{1}{4}\max\left(U\left[off_4\left(1,10,15000\right)\right],\left(\begin{array}{l}
\frac{1}{3}\max\left\{U\left[off_5\left(1,10\right)\right],EU\left(1,10\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(1,15000\right)\right],EU\left(1,15000\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(10,15000\right)\right],EU\left(10,15000\right)\right\}
\end{array}\right)\right)+\\
\frac{1}{4}\max\left(U\left[off_4\left(1,5000,15000\right)\right],\left(\begin{array}{l}
\frac{1}{3}\max\left\{U\left[off_5\left(1,5000\right)\right],EU\left(1,5000\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(1,15000\right)\right],EU\left(1,15000\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(5000,15000\right)\right],EU\left(5000,15000\right)\right\}
\end{array}\right)\right)+\\
\frac{1}{4}\max\left(U\left[off_4\left(10,5000,15000\right)\right],\left(\begin{array}{l}
\frac{1}{3}\max\left\{U\left[off_5\left(10,5000\right)\right],EU\left(10,5000\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(10,15000\right)\right],EU\left(10,15000\right)\right\}+\\
\frac{1}{3}\max\left\{U\left[off_5\left(5000,15000\right)\right],EU\left(5000,15000\right)\right\}
\end{array}\right)\right)
\end{array}\right)\right),
$$

CONTESTANTS' BELIEFS AND THE RATIONAL EXPECTATIONS HYPOTHESIS

In order to solve the decision problem, a forward-looking contestant needs to form expectations about the *Banker's offers* in subsequent rounds, so that she can figure out the possible consequences of her actions and maximize expected utility.

This implies that a researcher, without any prior information about the formulation of contestants' expectations, cannot identify contestants' preferences, because the distribution of preferences is not unique (Manski, 2004). For example, a contestant highly risk averse but strongly optimistic about future *Banker's offers* may behave, *ceteris paribus*, the same as a contestant highly risk loving but strongly pessimistic about future offers. Manski suggests asking people to self-report their expectations. Unfortunately, in a game like *Affari tuoi* this would imply asking contestants their beliefs about the future *Banker's offer* in dozens of possible lotteries.

To overcome this problem, researchers commonly rely on the hypothesis of people having rational expectations or, in other words, being capable of forecasting these consequences in a similar manner to a researcher with access to a statistical package.

All the other papers using "Deal or No Deal" data to infer contestants' risk attitude base the forward-looking problem on the hypothesis of rational expectations. Among the others:

- Andersen, S., G. Harrison, M. I. Lau, E. E. Rutström, 2006, "Dynamic choice behaviour in a natural experiment", University of Durham working paper in Economics and Finance
- De Roos, N., Sarafidis, Y., 2006, "Decision making under risk in Deal or No Deal", http://ssrn.com/abstract=881129

- Post, T., M. van Der Assem, G. Baltussen, R. Thaler, 2007, "Deal or No Deal? Decision making under risk in a large-payoff game show", *American Economic Review*, (possibly) forthcoming.

In contrast to the above papers, we recognise that the TV show *Affari tuoi* provides a suitable environment to estimate both preferences and beliefs without any prior assumptions about contestants' expectations, and to test if they predict according to rational expectations theory.

In effect, we recognise that in the 5^{th} round contestants' choices do not involve any expectations formation. The problem is just a simple choice between two lotteries: one with two equally probable prizes; the other being a degenerate lottery where the contestant can win the *Banker's offer* with probability equal to one.

So, using data on the 5^{th} round with those from choices in the 3^{rd} and 4^{th} round, under the hypothesis of the invariance over time of *contestants' risk attitude*, we are able to jointly estimate contestants' risk attitude and their *beliefs* about the *Banker's offers* in rounds 4 and 5 and also to test the hypothesis of rational expectations.

THE CHOICE PROCESS

We analyse the last 3 rounds of the game, which we denote by $n = 3,4,5$, under the hypothesis that players are expected utility maximizers.

We assume that player i's utility function takes the form of a Constant Absolute Risk Aversion (CARA) function, given by

$$U_i(x) = \frac{-\exp(-R_i \cdot x)}{R_i}, \tag{1}$$

where x is the outcome and R_i is the coefficient of absolute risk aversion for contestant i.

For identification, we normalize the utility function so that $U_i(0) = 0$ and $U_i(\max(x)) = 1$:

$$U_i(x) = \frac{1 - \exp(-R_i \cdot x)}{1 - \exp\left(-R_i \cdot \max(x)\right)}. \tag{2}$$

Let $U_i(off_n)$ be the utility of the *Banker's Offer* to contestant i in round n. Then, if contestant i exhibits EU preferences, she prefers the lottery over the *Banker's Offer* whenever $y_{in}^* > 0$, where y_{in}^* is defined by:

$$y_{in}^* = EU_{in}^{myopic}(X_{in}) - U_i(off_n) + \varepsilon_{in}, \tag{3}$$

if contestant i behaves myopically;

$$y_{in}^* = EU_{in}^{forward}(X_{in}) - U_i(off_n) + \varepsilon_{in}, \tag{4}$$

if contestant i is forward-looking.

In the two preceding formulae:

- $U_i(off_n)$ is the *Banker's offer* to contestant i in round n;
- $EU_{in}^{mypoic}(X_{in}) = \sum_{j=1}^{k_n} p_{jn} \cdot U_i(x_{jn})$ indicates the EU of the lottery in round n for player i, that is the probability-weighted utility of each outcome left in round n;
- $EU_{in}^{forward}(X_{in})$ is equal to:
  - ❖ $EU_{i3}^{forward} = \frac{1}{56}\sum_{j=1}^{56}\max\left\{U_i\left[off_4^e\left(X_{i4j}\right)\right]\ ,\ \frac{1}{10}\left[\sum_{k=1}^{10}\max\left[off_5^e\left(X_{i5jk}\right), EU\left(X_{i5jk}\right)\right]\right]\right\}$ in the 3rd round;

- ❖ $EU_{i4}^{forward} = \frac{1}{10}\left\{\sum_{j=1}^{10}\max\left[off_5^e\left(X_{i5j}\right), EU\left(X_{i5j}\right)\right]\right\}$ in the 4th round;
- ❖ $EU_{i5}^{forward} = EU\left(X_{i5}\right)$ in the 5th round

(where $EU(X_{in})$ is the probability-weighted utility of each outcome still available to contestant *i* in round *n*, and $off_n^e(.)$ is the offer contestant *i* expects to receive in round *n* (i.e. contestant *i*'s *beliefs* about the *Banker's offer* in round *n*). This is a known function of the lottery they will be confronted with.

- ε_{in} is a Fechner-type error term (Hey and Orme, *Econometrica*, 1994), with $\varepsilon_{in} \sim N(0, \sigma_\varepsilon^2)$. It has the interpretation of a computational error in calculating utilities. The variance of the error term is a measure of the magnitude of the error spread: the larger σ_ε^2, the greater the computational error.

Actually, we do not observe the continuous variable y^*_{in}, but we do observe the discrete variable $y_{in} = 1$ if individual i in round n chooses the lottery, and $y_{in} = 0$ if individual i in round n chooses the offer. The choice model is then described by:

$$\begin{aligned} y_{in} &= 1 \quad \text{if} \quad y^*_{in} > 0 \\ y_{in} &= 0 \quad \text{if} \quad y^*_{in} \leq 0 . \end{aligned} \tag{5}$$

As each game is composed of several binary choices, for each contestant we observe a sequence of only ones (if the contestant never accepts the money offer) or a sequence of ones followed by a zero (if the contestant accepts an offer). Then, the likelihood contribution of player i is the joint probability of observing the sequence of outcomes $(y_{i3},\ldots,y_{iN})$, where N is the round in which the game ends:

$$f\left(y_{i3},\cdots,y_{iN} \middle| X_{i3},\cdots,X_{iN}\right), \tag{6}$$

To handle this joint probability we need to make assumptions on the error term ε_{in} and on the independence of observations.

In effect, as our sample contains repeated observations on the same contestant, we cannot discard the hypothesis that these observations are correlated. In a linear random-effects panel data model, this situation is generally handled by introducing an individual-specific intercept in the model, referred to as unobserved heterogeneity, which is assumed to have a particular distribution across individuals. What is left of the error term is therefore independent of everything else in the model.

In contrast, our latent dependent variable is non-linear in the parameters to be estimated. In this case, to control for individual correlation we have at least two options:

1) assume that the unobserved heterogeneity is part of the error term (so that this component can be perceived as contestant *i* systematically overvaluing or undervaluing the difference between the expected utility of the lottery and the utility of the *Banker's offer*);
2) assume that there is a systematic individual-specific component in the risk aversion parameter that is normally distributed across the population:

$$R_i = \alpha + u_i^{ra}. \tag{7}$$

Here α a constant and u_i^{ra} reflects unobserved heterogeneity, with $u_i^{ra} \sim N(0, \sigma_u^2)$, such that $R_i \sim N(\alpha, \sigma_u^2)$.

After controlling for the unobserved heterogeneity in one of these two ways, we are allowed to assume that ε_{in} are independently and identically distributed, with $\varepsilon_{in} \sim \mathrm{N}(0, \sigma_\varepsilon^2)$, and independent of everything else. Our choice falls on the second option.

As we assume that all ε_{in} are independent over choices, in the case of EU preferences, we can write this joint probability as:

$$f\left(y_{i1},\ldots,y_{iN}\middle|X_i,\alpha\right)=\prod_{n=3}^{N}f\left(y_{in}\middle|X_i,\alpha\right)=$$
$$\int_{-\infty}^{+\infty}\left[\prod_{n=3}^{N}f(y_{in}|X_i,\alpha,u_i^{ra})\right]\frac{1}{\sigma_u}\phi\left(\frac{u_i^{ra}}{\sigma_u}\right)du_i^{ra}. \tag{8}$$

We also allow for the possibility of sub-optimal behaviour, by introducing a tremble parameter, ω (Loomes, Moffatt and Sugden , JRU, 2002). It measures the probability that contestants "tremble" (i.e. choose completely at random) in their choice. Taking this additional parameter into account, the last line of equation 8 becomes:

$$\int_{-\infty}^{+\infty}\left[\prod_{n=3}^{N}\left\{(1-\omega)f\left(y_{in}\middle|X_i,\alpha,u_i^{ra}\right)+\frac{\omega}{2}\right\}\right]\frac{1}{\sigma_u}\phi\left(\frac{u_i^{ra}}{\sigma_u}\right)du_i^{ra}. \tag{9}$$

To estimate the model we use the maximum simulated likelihood technique:

- the starting point of maximum likelihood is the assumption that the distribution of an observed phenomenon (the endogenous variable) is known, except for a finite number of unknown parameters. These parameters are estimated by taking those values for them that give the observed values the highest probability (likelihood);
- the probability in equation (20) (called likelihood contribution of individual *i*) is a complicated function of the parameters to be estimated. It can be solved by using Gauss-Hermite quadrature, but it requires a high computational time. This can be reduced by integrating over the unobserved heterogeneity by using Monte Carlo simulations and maximizing a simulated likelihood.

SOLUTION VIA SIMULATIONS

In 1981, Lerman and Manski introduced Monte Carlo simulation methods to approximate integrals of the following kind:

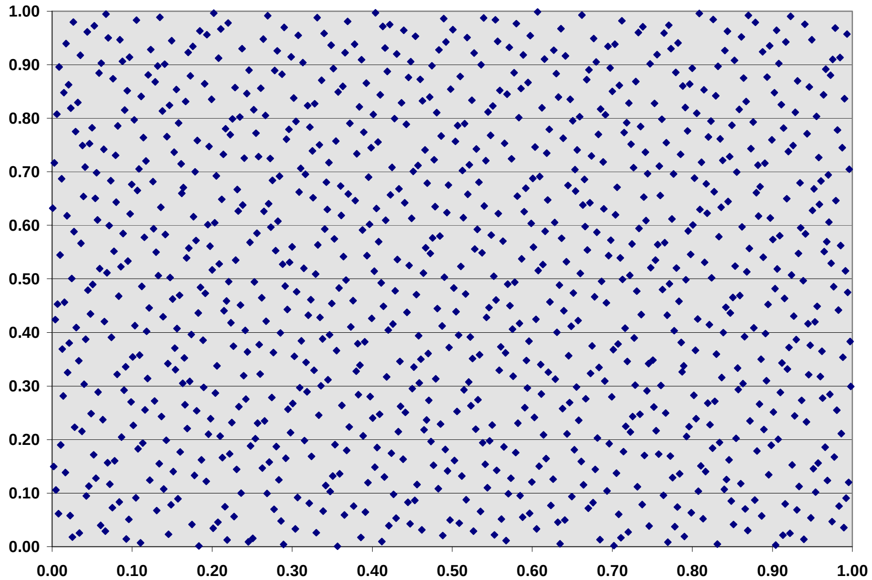
$$E[h(U)] = \int h(u)g(u)du\,, \tag{10}$$

where $U \sim g(u)$ and $h(u)$ is a function implied by the theory. They suggested that evaluating such an integral is asymptotically equivalent to computing a sample average of $h(u)$. Therefore, it is possible to draw R random variables from $g(u)$, say u^r, with r = 1,…,R, and to calculate

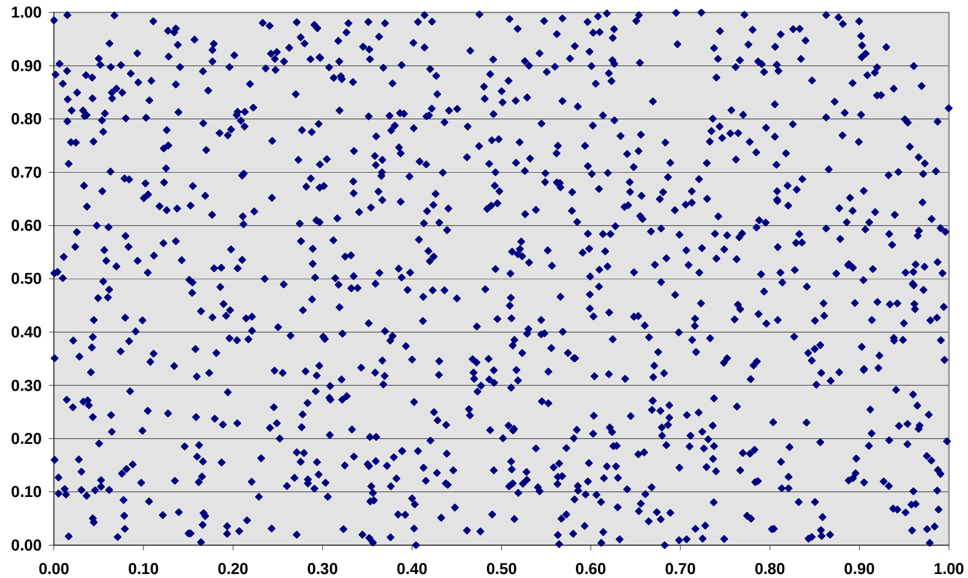
$$E_R[h(U)] = \frac{1}{R}\sum_{r=1}^{R} h(u^r) \tag{11}$$

in order to obtain an unbiased estimator of $E[h(U)]$, with a variance that goes to 0 as $R \to \infty$. This means that the higher the number of random draws we use, the more precise the estimator is, but this increases the computational time.

Anyhow, instead of using random draws, simulation can be based on *Halton Sequences*. These generate systematic quasi-random draws that, inducing a negative correlation over observations, allows using a small number of draws (many studies have shown that the results are more accurate with 100 Halton draws than 1000 random draws, that means a big gain in computational time). Halton sequences are based in primes and work as a "filling in the gaps" process. The figures below provide a comparison of the coverage of Halton draws and random draws for a bivariate uniform distribution.



Bivariate Halton draws based in primes 2 and 3



Bivariate random draws

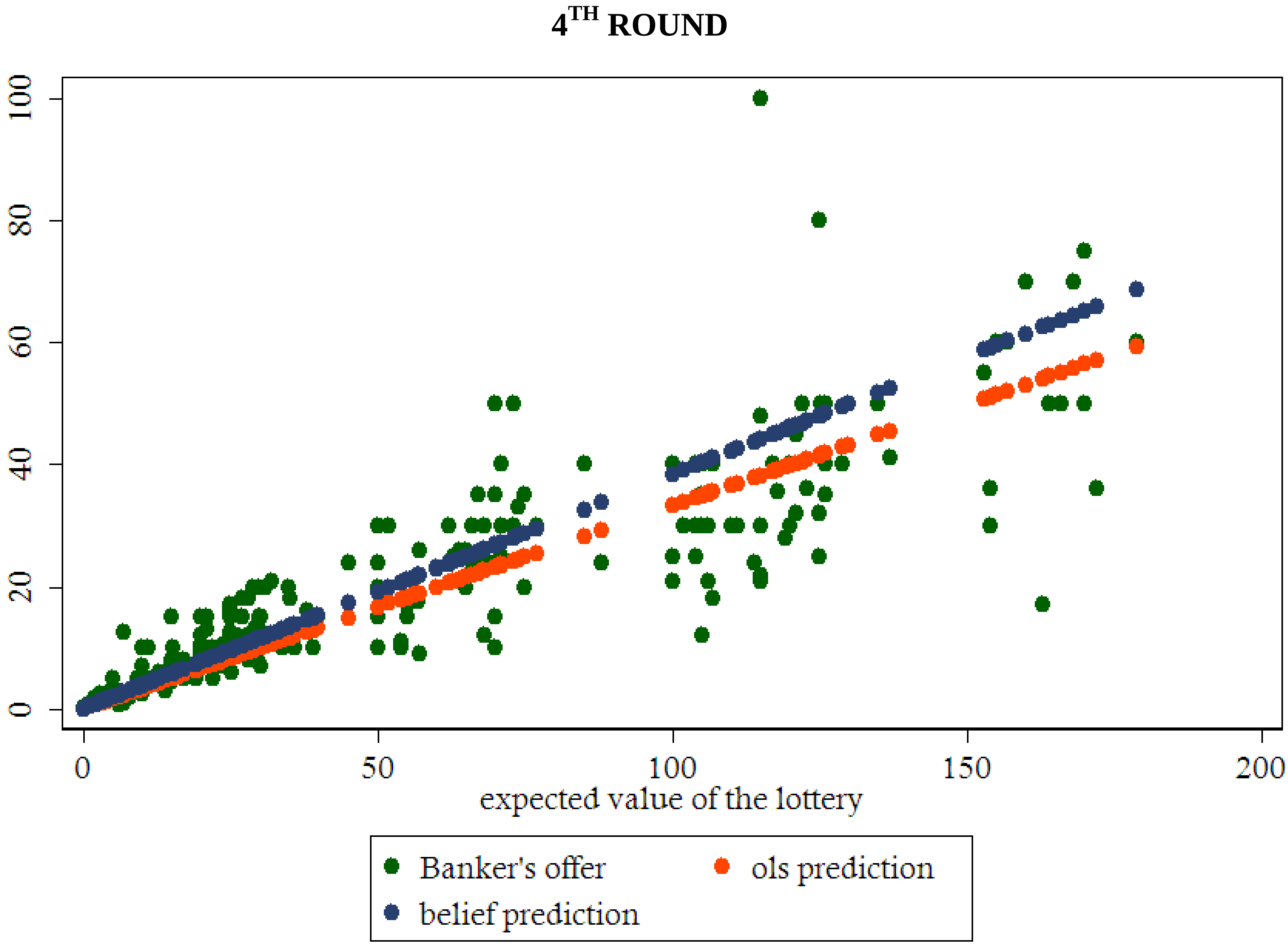
CHOICE MODEL ESTIMATES

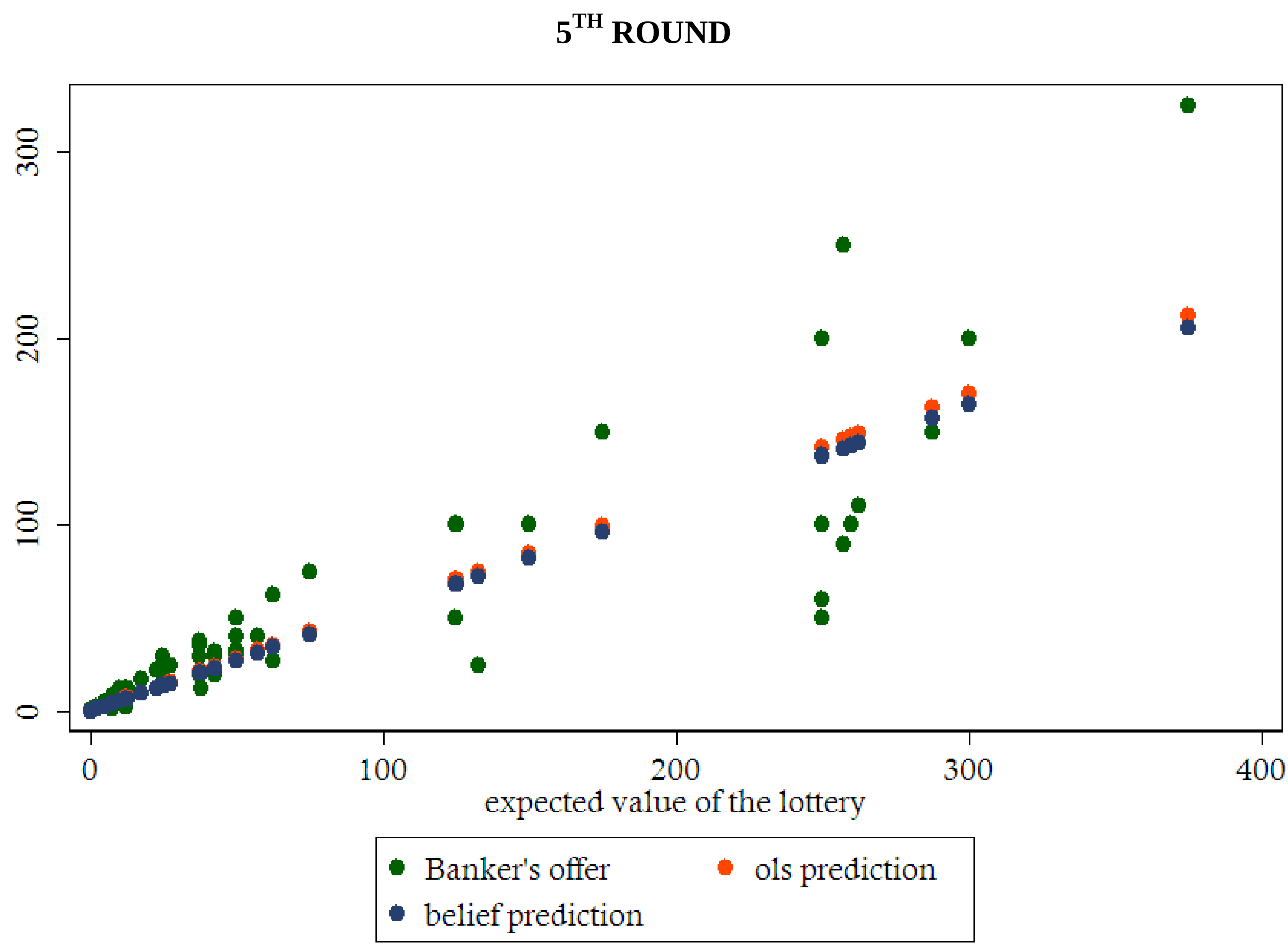
Parameter estimates; CARA utility functional; maximum simulated likelihood (294 contestants, 631 obs.)

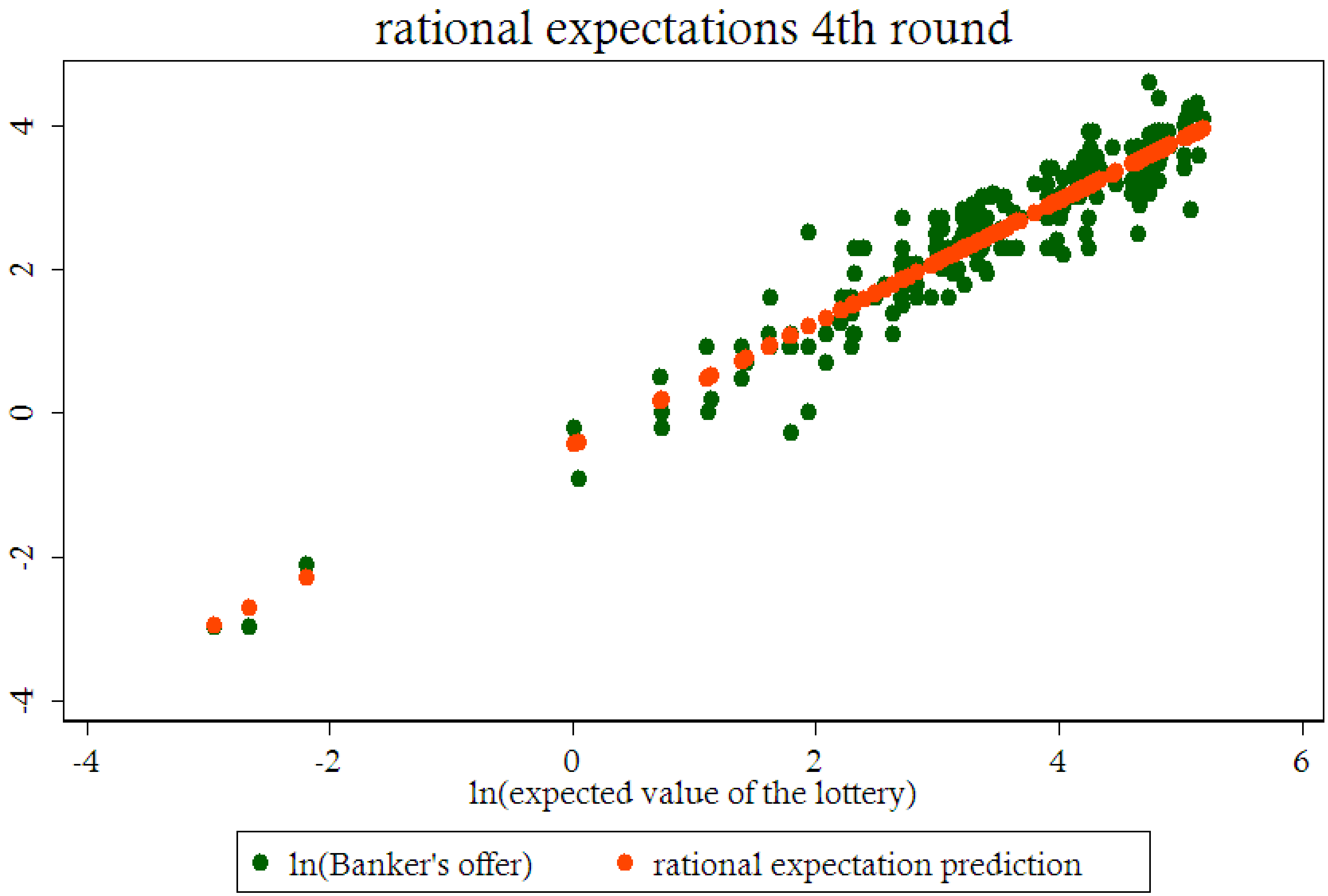
	Myopic contestants	***Forward-looking contestant***
Estimates of risk aversion, covariance matrix, scale parameters and tremble		
α	0.01236*** (0.00090)	0.02413*** (0.00217)
σ_u	0.00639*** (2.441e-06)	0.01779*** (3.7417578e-006)
σ_ε	0.03692*** (0.00008)	0.01311*** (0.00012)
ω	0.07275*** (0.02015)	0.06190** (0.02684)
Estimates of the belief equation for the 4th round $\beta(Exp.Val.Lott.)$		
β	-	0.38338*** (0.02896)
Estimates of the belief equation for the 5th round $\beta(Exp.Val.Lott.)$		
β	-	0.54776*** (0.04766)
Log-likelihood	-261.97520	-235.54324

**5% significance level

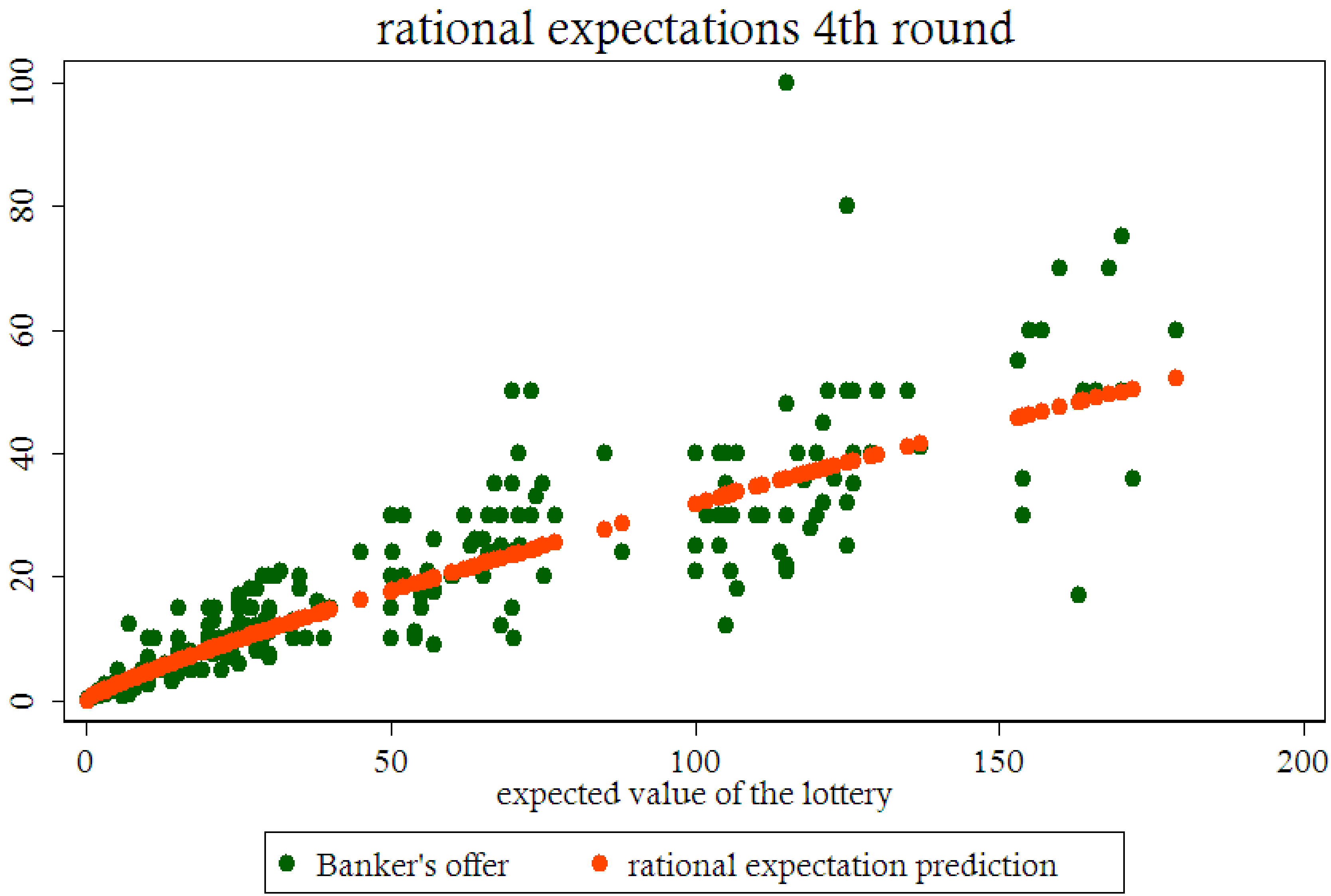
***1% significance level

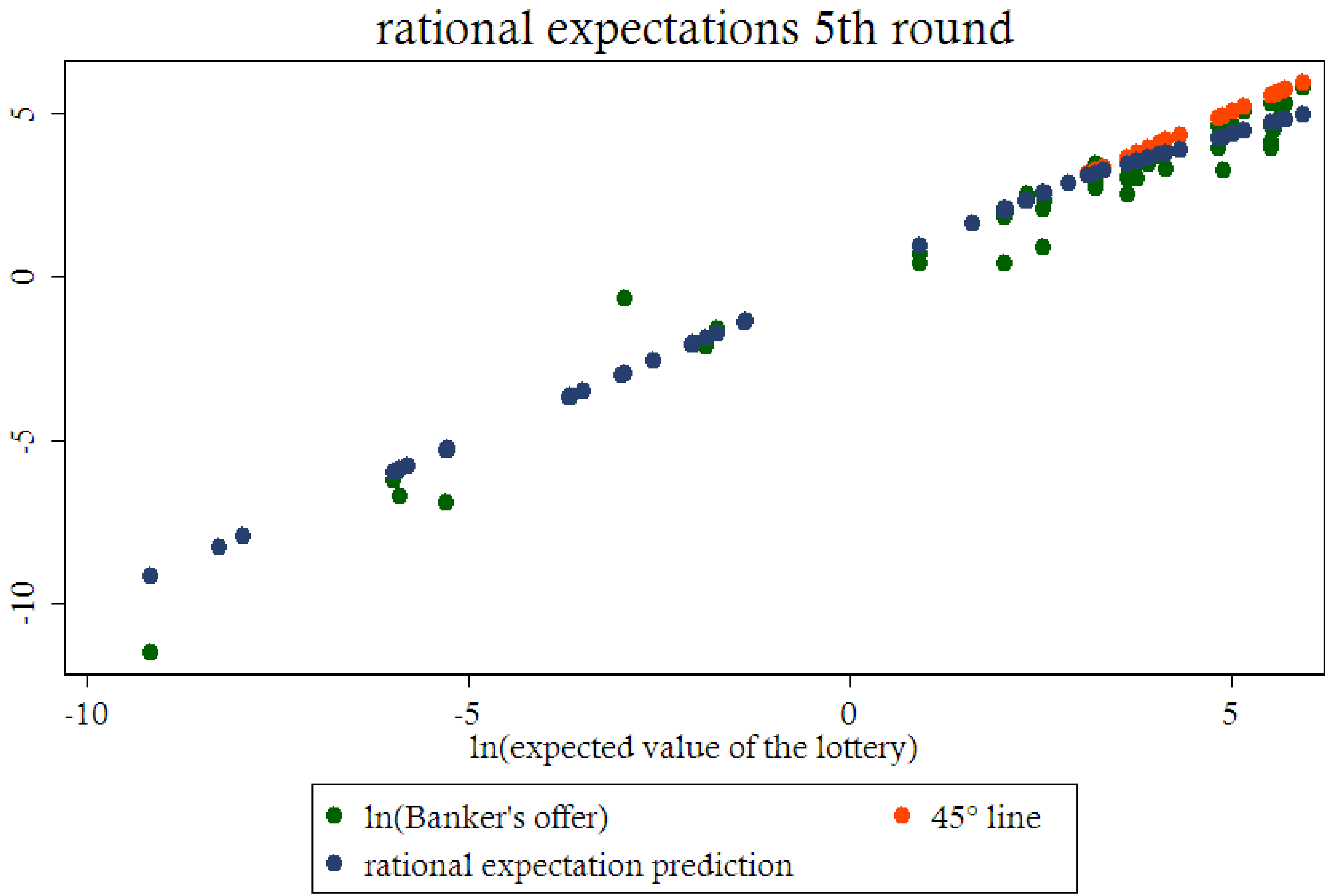

4TH ROUND
100
80
60
40
20
0
0
50
100
150
200
expected value of the lottery
Banker's offer
ols prediction
belief prediction


5TH ROUND
0
100
200
300
0
100
200
300
400
expected value of the lottery
Banker's offer
ols prediction
belief prediction



rational expectations 4th round
4
2
0
-2
-4
-4
-2
0
2
4
6
ln(expected value of the lottery)
ln(Banker's offer)
rational expectation prediction


rational expectations 4th round
0
20
40
60
80
100
0
50
100
150
200
expected value of the lottery
Banker's offer
rational expectation prediction


rational expectations 5th round
5
0
-5
-10
-10
-5
0
5
ln(expected value of the lottery)
ln(Banker's offer)
45° line
rational expectation prediction

CHOICE MODEL ESTIMATES UNDER THE HYPOTHESIS OF RATIONAL EXPECTATIONS

Parameter estimates; CARA utility functional; maximum simulated likelihood; (294 contestants, 631 obs.)

	Without constraints	***With constraints***
Estimates of risk aversion, covariance matrix, scale parameters and tremble		
α	0.02453*** (0.00358)	0.02524*** (0.00199)
σ_u	0.02004*** (9.5073592e-006)	0.02097*** (3.7417578e-006)
σ_ε	0.01225*** (9.4597598e-005)	0.02938*** (0.00011)
ω	0.04275*** (0.00055)	0.06190*** (0.00079)
Estimates of the belief equation for the 4th round $\exp(\gamma + \beta \ln(Exp.Val.Lott.))$		
γ	0.52476*** (0.14716)	-0.45500
β	0.64452 (0.03881)	0.85000
Estimates of the belief equation for the 5th round $\min[\ln(Exp.Val.Lott.)\ ,\ \exp(\gamma + \beta \ln(Exp.Val.Lott.))]$		
γ	-1.11559 (0.80626)	1.0000
β	1.12931*** (0.21191)	0.66667
Log-likelihood	-232.54020	-239.82445

***1% significance level

FUTURE WORKS

- mixture model of myopic and forward-looking contestants
- estimates of the two models using a rank-dependent expected utility functional