

National Centre of Competence in Research
Financial Valuation and Risk Management

Working Paper No. 106

**The α -Beauty Contest:
Choosing Numbers, Thinking Intervals**

Enrico De Giorgi Stefan Reimann

First version: June 2003
Current version: June 2007

This research has been carried out within the NCCR FINRISK project on
“Evolution and Foundations of Financial Markets”.

The α -Beauty Contest:

Choosing Numbers, Thinking Intervals*

Enrico De Giorgi[†]

Stefan Reimann[‡]

University of Lugano

University of Bielefeld

First Draft: June 14, 2003

This Version: June 8, 2007

*The authors are grateful to Francesco Audrino, Ernst Fehr, Thorsten Hens, Fabio Trojani, the editor Ehud Kalai, and two anonymous referees for valuable suggestions. We are grateful to Rosemarie Nagel for making data available. Financial support from the Foundation for Research and Development of the University of Lugano and from the National Centre of Competence in Research “Financial Valuation and Risk Management” (NCCR-FINRSIK) is gratefully acknowledged. The national centers of competence in research are managed by the Swiss National Science Foundation on behalf of the federal authorities. A preliminary version of this paper was presented at the 2004 International Meeting of the Economic Science Association, held in Amsterdam.

[†]Swiss Finance Institute and University of Lugano, via Buffi 13, CH-6900 Lugano, Switzerland, Tel. +41 +58 6664677, Fax. +41 +58 6664647, email: enrico.degiorgi@lu.unisi.ch (corresponding author).

[‡]Research Center Bielefeld-Bonn Stochastics, University of Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany, email: sreimann@physik.uni-bielefeld.de.

Abstract

We present a model for the α -beauty contest that explains common patterns in experimental data of one-shot and iterative games. The approach is based on two basic assumptions. First, players iteratively update their recent guesses. Second, players estimate intervals rather than exact numbers to cope with incomplete knowledge of other players' choices. Under these assumptions we extend the cognitive hierarchy model of Camerer, Ho, and Chong (2003a). The extended model is estimated on experimental data from a newspaper experiment.

Keywords: experiments, α -beauty contest, beliefs, cognitive hierarchy model.

JEL Classification Numbers: C70, C91, C93, D84.

SSRN Classification: behavioral finance, experimental studies.

1 Introduction

In his book *General Theory of Employment, Interest, and Money*, Keynes (1936, p. 156) argued that financial markets are like a beauty contest, where people guess which participant will be preferred by the others. While in real markets only few investor follows this strategy, Keynes' α -beauty contests are games that offer a simplified setting to study people's behavior.

Nagel, Bosch-Domènech, Satorra, and Garca-Montalvo (2002) succinctly describe the α -beauty contest (or guessing game): “Each player simultaneously chooses a decimal number in the interval $[0,100]$. The winner is the person whose number is closest to α times the mean of all chosen numbers, where $\alpha < 1$ is a predetermined and known number. The winner gains a fixed prize. If there is a tie, the prize is split amongst those who tie or a random draw decides the winner.”

The guessing game possesses a unique Nash equilibrium, which occurs when all players announce zero. Nevertheless, as Nagel (1995) showed in a number of experiments, the Nash equilibrium is generally not observed in the one-shot α -beauty contest; rather, it is approached after sufficiently many rounds in the iterative setting, i.e., in the presence of communication.

Field experiments were conducted to estimate the behavior of people when playing the guessing game in different settings. We refer mainly to studies with a larger number of players, for example the newspaper experiments conducted by Bosch-Domènech and Nagel (1997), Thaler (1997), and Selten and Nagel (1998), or the experiments with professional investors conducted by Montier (2004). Nagel (1995), however, showed that across different sample sizes, different

methods of collecting data, etc. the main patterns of people’s behavior remain stable. Nagel, Bosch-Domènech, Satorra, and Garca-Montalvo (2002) survey 24 experiments on one-shot guessing games. Figure 1 gives the results obtained from the newspaper experiment conducted by Bosch-Domènech and Nagel (1997).

[Figure 1 about here.]

Three main observations may be made from experimental data. The *first observation* is that the winning numbers observed in different (comparable) experiments are approximately equal. The *second observation* is that the spectrum of numbers announced is a superposition of a broad and highly skewed distribution and a collection of frequently chosen numbers $\{\chi_1, \chi_2, \dots\}$, which correspond to sequence $\{[\alpha^k (100/2)], k = 1, \dots\}$. We call these numbers “rational regimes”. While the skewed distribution reflects the players’ uncertainty about other players’ rationality, the rational regimes represent different depths of players’ thinking. The *third observation* is that equilibrium is established after a sufficient number of iterative rounds. Figure 2 reports the logarithm of the winning number as a function of the number of rounds played by the participants. We observe that the winning number converges to 0 as the number of rounds increases.

[Figure 2 about here.]

A reasonable model has to explain these facts consistently.

We start by considering the one-shot game setting. The strategy chosen by one individual crucially depends on her guess about the strategies of the others (first-order beliefs). Given

the basic setting, players have to choose a number from the interval $[0, 100]$. Assuming a large number of participants, if one player believes that all the others will announce the number 100, then her optimal reply is to announce $\alpha 100$.¹ This is the first thinking step. If the player further believes that the others will also perform the first thinking step, then she preferably responds by announcing the number $\alpha^2 100$. This is the second thinking step. In principle, the thinking process can be iterated infinitely. Some authors, including Nagel (1995), Stahl and Wilson (1995), Ho, Camerer, and Weigelt (1998), proposed the idea that players who perform k steps of thinking optimize (in terms of their best reply) against the $(k - 1)$ -step players. The cognitive hierarchy model of Camerer, Ho, and Chong (2003a, 2003b) assumes that all k -step players can accurately guess the relative proportion of players who are thinking $k' \leq k$ levels. More precisely, they assume that the number of k -step players is Poisson distributed with intensity τ . We refer to Camerer (2003) for an overview of these approaches. Models based on the idea that players execute thinking steps up to some level explain the rational regimes well, but fail to explain the background distribution which is observed in experimental data.

In most approaches, numbers – be they integers or reals – are considered as fundamental entities for individual choice of strategies. Of course it is not canonical that numbers have to be

¹As pointed out by one of the referees, in this example any number x with $|x - \alpha 100| < (1 - \alpha) 100$ will be closer to the winning number than 100 if the number of players is very large. Thus, if one player believes that all the others will announce 100, then she can choose any x satisfying the inequality above. However, in this paper, we defined the winning number as α times the mean of all chosen numbers and players' optimal replies are those minimizing the distance to the winning number.

chosen for computation. Other entities might be considered, especially if decision makers face uncertainty about the final outcome of their computation. In the guessing game, intervals rather than numbers might be chosen to cope with uncertain knowledge about the choice of the other participants. This fact is the key hypothesis of our model: *strategy choices rely on estimates on intervals rather than numbers*. The strategy of an individual player is the best reply to the choice of others under the assumption that the other players will choose their numbers in a interval around her own guess. For example, if a player guesses that the winning number is about 23, then the interval chosen in order to describe the choice of other participants might be $[7, 42]$. Without any further knowledge of the choices of other players and assuming that the player has minimal prejudice, she regards all numbers in this interval as equally probable.

We further assume that players perform thinking steps and the number of k -step players is Poisson distributed as in the cognitive hierarchy model of Camerer, Ho, and Chong (2003a). Our model is thus an extension of the latter and also reproduces the background distribution observed in experimental data.

We estimate our model on the results from the newspaper experiment conducted by Bosch-Domènech and Nagel (1997) and, similarly to previous works on the guessing game, we find that people perform on average approximatively 3.0 thinking steps. We also find strong evidence for our hypothesis that the numbers announced can be generated assuming that people estimate intervals rather than numbers. A high degree of uncertainty about other players' choices is obtained.

The remainder of the paper is organized as follows. In Section 2 we introduce the model for

the one-shot game. Section 3 presents estimation results from a newspaper experiment on the guessing game. Section 4 extends the model from the one-shot setting to the iterative setting. Section 5 concludes. Proofs and technical results are reported in the Appendix.

2 The model

There are $m = 1, \dots, M \leq \infty$ players, who choose a number x^m from $[0, N]$, where $N < \infty$.² They are supposed to guess the ‘winning number’ $y = \alpha \frac{1}{M} \sum_{m=1}^M x^m$, where $\alpha \in (0, 1)$ is fixed and known to all players. Let us consider a representative player m . For given choices $x^1, \dots, x^{m-1}, x^{m+1}, \dots, x^M$ of the other players $n \neq m$, the *optimal response* (or best reply) x^m of player m must satisfy $x^m = y$ and thus

$$x^m = \frac{\alpha}{M - \alpha} \sum_{n \neq m} x^n = \alpha \frac{M - 1}{M - \alpha} \bar{x}^{(-m)},$$

where $\bar{x}^{(-m)} = \frac{1}{M-1} \sum_{n \neq m} x^n$. Note that for M large enough, $\frac{M-1}{M-\alpha} \approx 1$ and thus $x^m \approx \alpha \bar{x}^{(-m)}$.

Since player m does not know in advance which numbers are chosen by the other players, she considers the choices of the others as random variables and builds her belief about their distributions. Let the random variable X^n denote the choice of player n in $[0, N]$. The random variable $\bar{X} = \frac{1}{M} \sum_{n=1}^M X^n$ is the average over all announced numbers and $Y = \alpha \bar{X}$ is the winning number. Moreover, $\bar{X}^{(-m)} = \frac{1}{M-1} \sum_{n \neq m} X^n$ and $Y^{(-m)} = \alpha \bar{X}^{(-m)}$. We make the following assumptions:

²The number of players M is usually known to all players. In newspaper experiments of the guessing game players can assume that M is very large.

Assumption 1. (i) For player m , the random variables $X^1, \dots, X^{m-1}, X^{m+1}, \dots, X^M$ are independent and identically distributed.

(ii) The number announced by player m corresponds to her optimal response to the expected choices of the other players, i.e.,

$$X^m = \alpha \frac{M-1}{M-\alpha} \bar{X}^{(-m)}.$$

Assumption 1 (i) is actually an assumption about the homogeneity of players and about the absence of communication between them.

Following the cognitive hierarchy model of Camerer, Ho, and Chong (2003a), we further assume that players use an *iterative thinking process* to guess the numbers which will be announced by other players. To indicate the depth of thinking, we add a subscript k to our previous notation. Assumption 1 is supposed to hold also for players' guesses after each thinking step.

Let us consider the thinking process of player m .

STEP 0 Player m assumes that all alternatives in $[0, N]$ are equally probable for all other players, i.e., she assumes that the X^n 's ($n \neq m$) are uniformly distributed in $[0, N]$. Under this assumption, $\bar{X}_0^{(-m)}$ is the average of $M-1$ independent uniformly distributed random variables on $[0, N]$ and

$$y_0^* = \mathbb{E}[Y] = \mathbb{E}[X_0^m] = \alpha \frac{M-1}{M-\alpha} \frac{N}{2}.$$

Note that y_0^* does not depend on m , i.e., it is the same for all players. This is an immediate consequence of Assumption 1.

Step 0 players believe that the other players randomly choose their numbers from $[0, N]$. Moreover, step 0 players assign the same probability to all numbers in $[0, N]$, i.e., they don't even assume that other players might have a minimal degree of rationality and avoid numbers like N that will never be a winning number. Indeed, some players seem to behave irrationally. For example, while one can argue that numbers strictly larger than αN are not winning numbers, Figure 1 shows that a significant proportion of players (the 4.5%) chose numbers larger than αN .

STEP 1 Player m takes into account the expected guessing number y_0^* she obtained from step 0. To include the possibility that her guess about the numbers chosen by the others is not exact, she considers an interval \mathcal{I}_1^m around this value. She assumes that given this interval the X_k^n 's are conditionally independent and uniformly distributed on \mathcal{I}_1^m for all $n \neq m$. More precisely, player m believes that for $n \neq m$ and conditioning on *unknown* realizations l_1^m, u_1^m of independent random variables $L_1^m \sim \text{unif}([0, y_0^*])$ and $U_1^m \sim \text{unif}([0, N - y_0^*])$ respectively, the interval is

$$\mathcal{I}_1^m = \mathcal{I}_1^m(l_1^m, u_1^m, \epsilon) = [y_0^* - \epsilon l_1^m, y_0^* + \epsilon u_1^m] \subset [0, N],$$

where $\epsilon \in [0, 1]$ is fixed. Parameter ϵ measures the players' degree of confidence in other players' rationality. When $\epsilon = 0$ then players are fully confident that other players will announce the expected winning number y_0^* from step 0. By contrast, if $\epsilon = 1$ step 1 players assume that the other players continue randomizing as in step 0. If $\epsilon < 1$, then the probability of numbers larger than $y_0^* + \epsilon(N - y_0^*)$ is zero. Thus, if step 1 players have some degree of confidence

about other players' rationality, some irrational choices will be excluded. For example, if $\epsilon < (\alpha N - \beta(N/2))/(N - \beta(N/2))$, then all numbers between αN and N are excluded by step 1 players.

The conditional expectation of X_1^m given l_1^m, u_1^m is

$$\mathbb{E}[X_1^m | (L_1^m, U_1^m) = (l_1^m, u_1^m)] = \alpha \frac{M-1}{M-\alpha} \left[y_0^* + \frac{\epsilon}{2}(u_1^m - l_1^m) \right].$$

Since l_1^m and u_1^m are unknown, player m will base her decision on the unconditional expectation

$$y_1^* = \mathbb{E}[X_1^m] = \alpha \frac{M-1}{M-\alpha} \left[y_0^* + \frac{\epsilon}{2} \left(\frac{N - y_0^*}{2} - \frac{y_0^*}{2} \right) \right] = \alpha \frac{M-1}{M-\alpha} \left[y_0^* \left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon N}{2} \right].$$

Again, y_1^* is the same for all players m .

The assumption about uniformly distributed intervals conditioned on uniformly distributed bounds is motivated by the argument of minimal knowledge. In other words, players' beliefs concerning other players' choices have a minimal structure. As discussed above, the choice of the confidence parameter is already an assumption on players' beliefs concerning other players' rationality. Thus, given their own degree of confidence, players do not impose further assumptions on other players' rationality and believe that step 1 players might have any interval around the step 0 unconditional expectation y_0^* and all intervals have the same probability of being chosen. As a consequence, the average interval is asymmetric around y_0^* , since obviously $y_0^* < N/2$.

It is clear that uniform distribution needs to be verified with experiments that capture the way players form their beliefs. However, at this initial stage, as long as the model produces a sequence of winning numbers that does not contradict the data, we prefer to use a simple model and "uniform distribution" seems to be the natural choice.

STEP k Given y_{k-1}^* from step $k-1$, player m believes that all players build intervals \mathcal{I}_k^m around y_{k-1}^* and the X_k^m 's are conditionally independent and uniformly distributed on \mathcal{I}_k^m . Analogously to Step 1, conditioning on *unknown* realizations l_k^m, u_k^m of independent random variables $L_k^m \sim \text{unif}([0, y_{k-1}^*])$ and $U_k^m \sim \text{unif}([0, N - y_{k-1}^*])$ respectively, the interval is given by

$$\mathcal{I}_k^m = \mathcal{I}_k^m(l_k^m, u_k^m, \epsilon) = [y_{k-1}^* - \epsilon l_k^m, y_{k-1}^* + \epsilon u_k^m] \subset [0, N],$$

The conditional expectation of X_k^m given l_k^m, u_k^m is

$$\mathbb{E}[X_k^m | (L_k^m, U_k^m) = (l_k^m, u_k^m)] = \alpha \frac{M-1}{M-\alpha} \left[y_{k-1}^* + \frac{\epsilon}{2}(u_k^m - l_k^m) \right],$$

and the unconditional expectation is

$$y_k^* = \mathbb{E}[X_k^m] = \alpha \frac{M-1}{M-\alpha} \left[y_{k-1}^* \left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon N}{2} \right].$$

The sequence of expected winning numbers $(y_k^*)_{k \geq 0}$ converges on a limit y^* as stated in the following Proposition.

Proposition 1. *Let $M > 0$, $N < \infty$, $\alpha \in [0, 1)$, and $(y_k^*, k \geq 0)$ the sequence of expected guessing numbers defined above. Then*

$$(1) \quad y_k^* = \beta^{k+1} \left(1 - \frac{\epsilon}{2}\right)^k \frac{N}{2} + \beta \frac{\epsilon N}{2} \frac{1 - \beta^k \left(1 - \frac{\epsilon}{2}\right)^k}{1 - \beta \left(1 - \frac{\epsilon}{2}\right)}.$$

where $\beta = \alpha \frac{M-1}{M-\alpha}$. Thus $(y_k^*)_{k \geq 0}$ is decreasing and $y_k^* \rightarrow y^*$ for $k \rightarrow \infty$, where

$$(2) \quad y^* = \frac{N}{2} \frac{\epsilon \beta}{2 - \beta(2 - \epsilon)}.$$

The proposition states that the thinking process generates a sequence of expected winning numbers that converges on an asymptotic expected winning number y^* , which depends on model parameters α , ϵ , N and the number of players M . The asymptotic winning number y^* is the fixed point of the iteration $y_k^* = \beta [y_{k-1}^*(1 - \epsilon/2) + (\epsilon/2)(N/2)]$ and thus solves $y^* = \beta [y^*(1 - \epsilon/2) + (\epsilon/2)(N/2)]$. Consequently, it is independent of the initial guess y_0^* . The convergence of the sequence (y_k^*) is very fast, so that y^* well approximates the expected winning numbers obtained after only a few levels of thinking. Remember that Nagel (1995) conjectured that individuals use only approximately three thinking steps. The convergence speed in our model is shown in Table 1, where we report the sequence of winning numbers generated by the thinking process for several values of ϵ and α . We also observe that the convergence speed increases with ϵ . In other words, if players are less confident about their guesses, their uncertainty over other players' choices will reduce the effect of additional thinking steps, so that players may choose not to update their guesses further. Also the number of players M has an impact on the sequence of unconditional expectations (y_k^*) and the convergence speed. Indeed, both are governed by parameter β , which varies from $(\alpha)/(1 - \alpha)$ if there are only two players (for $\alpha = 2/3$ it corresponds to $1/2$) to α if the number of players M is very large. Kocher and Sutter (2005) found that the size of the population has an impact on the winning number. Our model is consistent with this observation, even though the impact of M on the sequence of unconditional expectations (y_k^*) is not very large. Nevertheless, in our model M has a stronger impact on the distribution of announced number, as will become clear in Proposition 2 below.

From Proposition 1 we also obtain that y^* is strictly positive, unless $\epsilon = 0$, i.e., the asymp-

otic expected winning number differs from the game-theoretic equilibrium 0 if the confidence parameter is strictly positive.

[Table 1 about here.]

Next, we derive the probability density function of the winning number at each level of thinking and asymptotically, i.e., for $k \rightarrow \infty$. Owing to the rapid convergence of the thinking process on the asymptotic value, the asymptotic distribution is a good approximation for the distribution of winning numbers even after 3-4 thinking steps.

Proposition 2. *Let $M > 1$, $N < \infty$, and $\alpha \in (0, 1)$. Let $\epsilon > 0$, then the unconditional density function f_k of X_k^m for $k > 1$ is given by³*

$$f_k(x; \epsilon) = \frac{1}{y_{k-1}^* (N - y_{k-1}^*)} \int_0^{N - y_{k-1}^*} \int_0^{y_{k-1}^*} g(x; l, u, y_{k-1}^*, \epsilon) dl du,$$

where

$$g(x; l, u, y, \epsilon) = \frac{M - 1}{2\beta\epsilon(M - 2)!} \sum_{r=0}^{M-1} (-1)^r \binom{M - 1}{r} \frac{(h(x; l, u, y, \epsilon) - r)^{M-2} \text{sign}(h(x; l, u, y, \epsilon) - r)}{u + l}$$

and

$$h(x; l, u, y, \epsilon) = (M - 1) \frac{x - \beta(y - \epsilon l)}{\beta\epsilon(u + l)}.$$

The density function f_k gives the distribution of the winning number after k levels of the thinking process. It explains the dispersion of announced numbers around the unconditional

³We explicitly indicate that the unconditional density function also depends on ϵ because in the next section we will assume that ϵ is randomly chosen from a given set. In this case $f_k(\cdot; \epsilon)$ is the density function conditioning on the realization of ϵ .

expectations. By contrast, the background distribution observed in real data is not captured by models that exploit only the idea of different thinking levels. Figure 3 shows the density functions f_k for $k = 1, 2, 5$ and several values of ϵ . For larger values of ϵ the density functions are more dispersed, reflecting the small degree of confidence about other players' choices. If ϵ is close to zero, the density function is concentrated around the y_k^* 's.

[Figure 3 about here.]

The following corollary is an immediate consequence of the previous proposition and gives the unconditional density function of the asymptotic winning number:

Corollary 1. *Let $M > 1$ and $\alpha \in [0, 1)$. Let $\epsilon > 0$ then the unconditional density function f_∞ of $X_\infty^m = \lim_{k \rightarrow \infty} X_k^m$ is given by*

$$f_\infty(x; \epsilon) = \frac{1}{y^* (N - y^*)} \int_0^{y^*} \int_0^{N-y^*} g(x; l, u, y^*, \epsilon) dl du,$$

where g is defined in Proposition 2.

Density function f is illustrated in Figure 4 for several values of M . In Appendix B we apply the Central Limit Theorem to derive an approximation of the density functions f_k and f_∞ if the number of players M is large. For comparison, Figure 4 also reports the approximated asymptotic density function.

[Figure 4 about here.]

3 Maximum-likelihood estimation

In this section we estimate our model for the guessing game on the results from a newspaper experiment with 3696 participants conducted by Bosch-Domènech and Nagel (1997). The histogram of the numbers announced is shown in Figure 1. Approximately one-quarter of the players have announced numbers which correspond to rational regimes or belong to 5%-intervals around rational regimes. The remaining three-quarters of players have announced numbers which cannot be approximated by a rational regime. If the number of players is very high, our model does not explain announced numbers which are higher than $N/2$.⁴ Thus, we eliminate those numbers from our data set and we obtain 3352 observations.⁵

Our results from Section 2 are obtained under the assumption that all players perform the same number of thinking levels and maintain the same confidence parameter ϵ over the sequence of thinking steps. The advantage of these assumptions is that we are able to derive a simple

⁴In our model any number in $[0, N]$ can be generated from step 0 players beliefs on other players' choices. Nevertheless, we assume that each player announces her number according to her best reply to other players' choices. Therefore, the number announced will be higher than $N/2$ only if one player believes that the majority of the other players are irrational step 0 players. However, this case is not captured in our model by Assumption 1 and the uniform distribution assumption, which implies that the probability that all other players are irrational is zero. By contrast, if the number of player M is small, then other players choices have larger impact on players' replies, and therefore our model can explain also numbers that are larger than $N/2$.

⁵This means that our model captures 91% of players' choices in the newspaper experiment discussed in this section. By contrast, the cognitive hierarchy model of Camerer, Ho, and Chong (2003a) only captures the rational regimes, thus approximately 25% of players' choices.

closed form solution for the expected winning number at each level of thinking. However, the experimental results (see Figure 1) do not support the assumption of a fixed number of thinking steps for all players, and there is no evidence that the confidence level is constant over the thinking process. Moreover, players' confidence parameters might substantially differ. For example, the skew background distribution of announced numbers observed in experiments with a large number of players suggests that some players exhibit a low degree of confidence, i.e., ϵ is close to 1. By contrast, we also observe that a large number of players choose the rational regimes, which suggest that many players might have a high degree of confidence, thus ϵ close to 0.

Therefore, in order to estimate the model, in this section we relax the assumptions about the number of thinking levels k and the confidence parameter ϵ . For the thinking levels we adopt the cognitive hierarchy model of Camerer, Ho, and Chong (2003a) and we assume a Poisson distribution with parameter τ . For the parameter ϵ we impose a minimal structure. Consistently with the observation that a high proportion of players choose the rational regimes, we allow many players to have ϵ close to zero. Thus, we assume that ϵ is distributed on the set $\mathcal{E} = \{2^{-i} : i = 0, \dots, I\}$ with (unknown) probabilities $p_i = \mathbb{P}[\epsilon = 2^{-i}]$ for $i = 1, \dots, I$.⁶ Under these assumptions for the number of thinking levels and the degree of confidence, our model is

⁶We can also assume that set \mathcal{E} depends on the thinking levels. Indeed, in Figure 1 we observe that the distributions around the first rational regimes are less dispersed, suggesting that ϵ might be smaller in the first steps of the thinking process. However, in order to justify more complex assumptions on the distribution of ϵ further research would be required.

an extension of the cognitive hierarchy model of Camerer, Ho, and Chong (2003a). Indeed, in this latter model it is assumed that $\epsilon = 0$ for all players.

We estimate the parameters of our model using maximum-likelihood estimation. From Proposition 2 we know the density functions f_k of the winning numbers after each thinking level, so that given our assumptions on ϵ and k , the density function for the announced numbers can be easily obtained and corresponds to:

$$f(x) = \sum_{i=0}^I p_i \sum_{k=0}^{\infty} f_k(x; 2^{-i}) \frac{\tau^k}{k!} e^{-\tau}.$$

We see from Proposition 2 that the density functions f_k are difficult to handle analytically, particularly if the number of players M is large, which is the case we focus on in this section. However, we show in Appendix B that if M is large, we can use the Central Limit Theorem to approximate f_k by a density function \tilde{f}_k which is analytically more tractable. We rely on the approximated density functions in order to derive the maximum-likelihood estimation of parameter τ and the distribution of ϵ on $[0, 1]$.

We make two different assumptions on the distribution of ϵ on \mathcal{E} . First, we assume that ϵ is uniformly distributed on \mathcal{E} (i.e., $p_i = 1/(I + 1)$ for $i = 0, \dots, I$) and estimate I from the data. Second, we fix the parameter I equal to the estimated parameter from the previous model, and we estimate τ and the probabilities $p_1, \dots, p_{I-1} \in [0, 1]$ from the data. Obviously, the second model is an extension the first one. Table 2 reports the results and Figure 5 shows the estimated distributions of the announced numbers under the two different assumptions on ϵ .

[Table 2 about here.]

Our estimation results confirm previous works on the guessing game which found for the same data set an average number of thinking levels equal to 3. The estimated probabilities for the confidence parameter ϵ confirms that a large number of players are quite confident about other players' rationality (the probability that ϵ is smaller or equal than 0.25 is 48.3%). We also find that in order to explain the distribution of the announced numbers we need to assume that some players are less confident about other players' rationality and the probability that players have $\epsilon = 1$ is quite large. Nevertheless, these results should not be over-emphasized, because some assumptions have to be tested with a larger set of experimental results. Indeed, more experiments on the guessing games are needed in order to build a parsimonious model for the confidence parameter ϵ that captures not only the distribution of ϵ on $[0, 1]$, but also the way players update their confidence parameter over thinking steps.

[Figure 5 about here.]

4 Convergence towards equilibrium

We finally consider the case that the guessing game is played for a number of rounds. The rules are the same for any round and are identical to the one-shot game discussed previously. Between any two rounds the winning number determined in the preceding round is made public and all players update their beliefs before performing the next round. We assume that all players perform an arbitrarily large number of thinking iterations within each round. This assumption is motivated by the rapid convergence of the thinking process within each round. Therefore,

we assume that the winning number announced at the end of each period is the respective asymptotic winning number (see Proposition 1) of this round, denoted by $y^*(t)$.

The asymptotic winning number y^* of the one-shot game can also be written as

$$y^* = y_0^*(N) \frac{\epsilon}{2 - \beta(2 - \epsilon)},$$

where $y_0^*(N) = \beta(N/2)$ corresponds to the initial expected winning number if the interval of possible winning numbers is $[0, N]$. We recall that the parameter β corresponds to $\alpha(M - 1)/(M - \alpha)$. We define function $c(\cdot)$ by

$$c(\epsilon) = \frac{\epsilon}{2 - \beta(2 - \epsilon)}.$$

Thus the asymptotic expected winning number corresponds to the expected winning number $y_0^*(N)$ on $[0, N]$ multiplied by a factor $c(\epsilon) < 1$ depending only on the confidence parameter ϵ (here β and M are given by the game setup). Therefore, if players were informed of the initial expected winning number y_0^* without any knowledge of the length N of the interval they would infer that $N = (2/\beta) y_0^*$ and their thinking process would generate the asymptotic expected winning number $y^* = y_0^* c(\epsilon)$. This observation is the starting point of our discussion on the iterated guessing game.

According to the one-shot game, the winning number announced at the end of round 1 is $y^*(1) = c(\epsilon_1) y^*(0)$, where ϵ_1 is the confidence parameter of the first round and $y^*(0) = \beta(N/2)$. The number $y^*(1)$ is made public to all players at the beginning of round 2, so that all players update their beliefs by restricting their initial interval $[0, N]$ to a smaller interval $[0, (2/\beta) y^*(1)]$. They then start their new thinking process of round 2 as the one-shot game is played on the

smaller interval and arrive at $y^*(2) = y^*(1) c(\epsilon_2)$. Note that since parameters β and M are given by the rule of the game, $y^*(t)$ essentially depends only on the sequence $(\epsilon_1, \dots, \epsilon_t)$ of confidence parameters up to time t . It might be expected that because of some adaptive mechanism, ϵ_t might change over rounds. As a first order approximation we assume that the confidence parameter is constant over rounds, i.e., $\epsilon_t = \epsilon$ for all t . Under these assumptions the iterative setting is governed by the following recurrence equation

$$(3) \quad y^*(t) \approx y^*(t-1) \frac{\epsilon}{2 - \beta(2 - \epsilon)},$$

where $y^*(0) = \beta(N/2)$. Hence

$$y^*(t) \approx y^*(0) \left(\frac{\epsilon}{2 - \beta(2 - \epsilon)} \right)^t,$$

or

$$\ln y^*(t) \approx \ln y^*(0) + t \ln \left(\frac{\epsilon}{2 - \beta(2 - \epsilon)} \right).$$

Therefore, since $\epsilon/(2 - \beta(2 - \epsilon)) < 1$, as $t \rightarrow \infty$ the expected winning number $y^*(t)$ converges on 0. The convergence speed depends on ϵ , α , and the players M . We refer to Figure 2, which reports experimental results from the iterative guessing game. The log-linear relationship fits real data very well. The same result applies to the other data set that we analyzed. The parameter ϵ is estimated on the results of different experiments on the iterative game using linear regression. The estimated values are reported in Table 3. It is shown that ϵ is consistently larger than 0. Moreover, we also observe that ϵ varies from experiment to experiment, supporting our hypothesis that people show broad heterogeneity with respect to their degree of confidence. We

also point out that the number of players in the experiment of the iterative setting is usually small.

[Table 3 about here.]

5 Conclusion

In this paper, we studied the α -beauty contest with $\alpha \in (0, 1)$ and in different settings, such as the one-shot setting and the iterative setting with communication between rounds. Our considerations started from investigating real data. Experimental data from the one-shot setting of the 0-equilibrium games exhibit a common pattern: the spectrum of announced numbers is a superposition of a skewed continuous background distribution and a regime of frequently chosen numbers. Moreover, in the one-shot setting the unique Nash equilibrium 0 is usually not observed, while it is established in the iterative setting after a sufficient number of rounds.

Our model is an extension of the cognitive hierarchy model of Camerer, Ho, and Chong (2003a) that is able to explain this pattern. It is based on the assumption that players successively update their recent beliefs by estimating intervals rather than numbers. The model has one free parameter ϵ , which is a measure of the confidence of players. It was shown analytically that if players have only finite confidence, the expected winning number is strictly positive. The Nash equilibrium is obtained if players' confidence is infinite and they perform a large number of thinking steps.

Our estimation results show that on average players only execute three thinking levels and

possess heterogenous degree of confidence. As a consequence, our model produce a highly skewed distribution of announced numbers in agreement with real data. Moreover, in our model, a straightforward consequence of the fast convergence over thinking steps is that convergence towards equilibrium in the iterative setting of the game is a geometric series in the number of rounds played and the winning number over rounds converges on zero.

References

- BOSCH-DOMÈNECH, A., AND R. NAGEL (1997): “El juego de adivinar el número X: una explicación y la proclamacion del vencedor,” *Espansión*, June 16, 42–43.
- CAMERER, C. (2003): *Behavioral Game Theory: Experiments on Strategic Interaction*. Princeton University Press, Princeton.
- CAMERER, C., T. HO, AND J. CHONG (2003a): “A Cognitive Hierarchy Model of One-Shot Games,” *Quarterly Journal of Economics*, 119, 861–898.
- CAMERER, C., T. HO, AND K. CHONG (2003b): “Models of Thinking, Learning, and Teaching in Games,” *American Economic Review*, 93(2), 192–195.
- HO, T., C. CAMERER, AND K. WEIGELT (1998): “Iterated Dominance and Iterated Best Response in Experimental “p-Beauty Contests”,” *American Economic Review*, 88(4), 947–969.

- KEYNES, J. M. (1936): *The General Theory of Interest, Employment and Money*. Macmillan, London.
- KOCHER, M. G., AND M. SUTTER (2005): “The Decision Maker Matters: The Individual Versus Group Behaviour in Experimental Beauty-Contest Games,” *Economic Journal*, 115(500), 200–223.
- MONTIER, J. (2004): “Who’s a Pretty Boy Then? Or Beauty Contests, Rationality and Greater Fools,” Dresdner Kleinwort Wasserstein – Global Equity Strategy, Available at SSRN: <http://ssrn.com/abstract=505902>.
- NAGEL, R. (1995): “Unraveling in Guessing Games: An Experimental Study,” *American Economic Review*, 85(5), 1313–1326.
- NAGEL, R., A. BOSCH-DOMÈNECH, A. SATORRA, AND J. GARCA-MONTALVO (2002): “One, Two, (Three), Infinity: Newspaper and Lab Beauty-Contest Experiments,” *American Economic Review*, 92(5), 1687–1701.
- SELTEN, R., AND R. NAGEL (1998): “Das Zahlenwahlspiel – Hintegrunde und Ergebnisse,” *Spektrum der Wissenschaft*, February, 16–22.
- STAHL, D. O., AND P. WILSON (1995): “On Players Models of Other Players: Theory and Experimental Evidence,” *Games and Economic Behavior*, 10(1), 218–254.
- THALER, R. (1997): “Giving Markets a Human Dimension,” *Financial Times*, June 16.

A Proofs

A.1 Proof of Proposition 1

The sequence y_k^* is defined as

$$y_0^* = \beta \frac{N}{2} \quad \text{and} \quad y_k^* = \beta \left[y_{k-1}^* \left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon}{2} \frac{N}{2} \right] \quad \text{for } k \geq 1.$$

By applying iteratively the last equation we obtain

$$\begin{aligned} y_k^* &= \beta \left\{ \beta \left(y_{k-2}^* \left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon}{2} \frac{N}{2} \right) \left(1 - \frac{\epsilon}{2}\right) + \frac{\epsilon}{2} \frac{N}{2} \right\} \\ &= \beta^2 \left(1 - \frac{\epsilon}{2}\right)^2 y_{k-2}^* + \alpha^2 \frac{\epsilon}{2} \frac{N}{2} \left(1 - \frac{\epsilon}{2}\right) + \beta \frac{\epsilon}{2} \frac{N}{2} \\ &= \dots = \beta^k \left(1 - \frac{\epsilon}{2}\right)^k y_0^* + \beta \frac{\epsilon}{2} \frac{N}{2} \sum_{l=0}^{k-1} \beta^l \left(1 - \frac{\epsilon}{2}\right)^l \\ &= \beta^{k+1} \left(1 - \frac{\epsilon}{2}\right)^k \frac{N}{2} + \beta \frac{\epsilon}{2} \frac{N}{2} \frac{1 - \beta^k \left(1 - \frac{\epsilon}{2}\right)^k}{1 - \beta \left(1 - \frac{\epsilon}{2}\right)}. \end{aligned}$$

For $k \rightarrow \infty$ and $\beta \neq 1$ ($\beta < 1$ if and only if $\alpha < 1$) we obtain

$$y_k^* \rightarrow y^* = \frac{N}{2} \frac{\epsilon \beta}{2 - \beta(2 - \epsilon)}.$$

A.2 Proof of Proposition 2

For $n \neq m$, let $\tilde{X}_k^n = \frac{X_k^n - (y_{k-1}^* - \epsilon l_k^m)}{\epsilon (u_k^m + l_k^m)}$. Then, conditionally on the lower and upper bounds l_k^m, u_k^m for the interval $\mathcal{I}_k^m(l_k^m, u_k^m, \epsilon)$, the random variables \tilde{X}_k^n are independent, identically uniformly

distributed on $[0, 1]$. We have:

$$\begin{aligned}\mathbb{P}[X_k^m \leq x | l_k^m, u_k^m] &= \mathbb{P}\left[\beta \bar{X}_k^{(-m)} \leq x | l_k^m, u_k^m\right] \\ &= \mathbb{P}\left[\sum_{n \neq m} \tilde{X}_k^n \leq (M-1) \frac{x - \beta(y_{k-1}^* - \epsilon l_k^m)}{\beta \epsilon (u_k^m + l_k^m)} | l_k^m, u_k^m\right]\end{aligned}$$

Let $h(x; l, u, y, \epsilon) = (M-1) \frac{x - \beta(y - \epsilon l)}{\beta \epsilon (u + l)}$. Then, $\frac{\partial}{\partial x} h(x; l, u, y, \epsilon) = \frac{M-1}{\beta \epsilon (u+l)}$. Since conditionally on l_k^m and u_k^m , $\sum_{n \neq m} \tilde{X}_k^n$ is the sum of $M-1$ independent, identically and uniformly distributed random variables on $[0, 1]$ it follows:

$$\begin{aligned}g(x; l_k^m, u_k^m, y_{k-1}^*, \epsilon) &= \frac{d}{dx} \mathbb{P}[X_k^m \leq x | l_k^m, u_k^m] \\ &= \frac{M-1}{\beta \epsilon (u_k^m + l_k^m)} \\ &\quad \times \frac{1}{2(M-2)!} \sum_{r=0}^{M-1} (-1)^r \binom{M-1}{r} (h(x; l_k^m, u_k^m, y_{k-1}^*) - r)^{M-2} \text{sign}(h(x; l_k^m, u_k^m, y_{k-1}^*) - r).\end{aligned}$$

Finally, l_k^m and u_k^m are the realizations of independent, uniformly distributed random variables on $[0, y_{k-1}^*]$ and $[0, N - y_{k-1}^*]$, respectively the statement of the proposition follows.

B Maximum-likelihood estimation

For $k > 1$, we have that $X_k^m = \beta \frac{1}{M-1} \sum_{n \neq m} X_k^n = \frac{1}{M-1} \sum_{n \neq m} \beta X_k^n$. Let $\tilde{X}_k^m = \beta X_k^n$, then conditioning on l_k^m and u_k^m the random variables X_k^n , $n \neq m$, are independent, identically and uniformly distributed on $[\beta y_{k-1}^* - \epsilon \beta l_k^m, \beta y_{k-1}^* + \epsilon \beta u_k^m]$. We compute the conditional mean

$m_k(l_k^m, u_k^m; \epsilon)$ and the conditional variance $s_k^2(l_k^m, u_k^m; \epsilon)$ of X_k^n , given l_k^m and u_k^m :

$$\begin{aligned} m_k(l_k^m, u_k^m; \epsilon) &= \beta (y_{k-1}^* + \frac{1}{2} \epsilon (u_k - l_k^m)) \\ s_k^2(l_k^m, u_k^m; \epsilon) &= \frac{1}{12} \epsilon^2 \beta^2 (u_k(l_k^m, u_k^m; \epsilon) + l_k(l_k^m, u_k^m; \epsilon))^2. \end{aligned}$$

By the central limit theorem, if $M \rightarrow \infty$

$$\frac{X_k^m - m_k(l_k^m, u_k^m; \epsilon)}{\frac{s_k(l_k^m, u_k^m; \epsilon)}{\sqrt{M-1}}} \rightarrow N(0, 1).$$

Therefore, for M large enough and conditionally on l_k^m and u_k^m , each random variable X_k^m is approximatively $N(m_k(l_k^m, u_k^m; \epsilon), \frac{s_k^2(l_k^m, u_k^m; \epsilon)}{M-1})$ and

$$f_k(x; \epsilon) \approx \tilde{f}(x; \epsilon) = \frac{1}{y_k^*} \frac{1}{N - y_k^*} \int_0^{y_{k-1}^*} \int_0^{N - y_{k-1}^*} \frac{\sqrt{M-1}}{s_k(l, u; \epsilon) \sqrt{2\pi}} e^{-\frac{1}{2}(M-1) \frac{(x - m_k(l, u; \epsilon))^2}{s_k^2(l, u; \epsilon)}} du dl.$$

Similarly, for $k = 0$:

$$f_0(x) \approx \tilde{f}_0(x) = \frac{\sqrt{M-1}}{s_0 \sqrt{2\pi}} e^{-\frac{1}{2}(M-1) \frac{(x - m_0)^2}{s_0^2}}$$

where $m_0 = \beta \frac{N}{2}$ and $s_0^2 = \frac{1}{12} \beta^2 N^2$.

Let K be the number of thinking levels. We assume that $K \sim \text{Poisson}(\tau)$, i.e., $\mathbb{P}[K = k] = \frac{\tau^k}{k!} e^{-\tau}$. Moreover, $\epsilon \in \{2^{-i}, i = 0, \dots, I\}$, where $I > 0$ is such that 2^{-I} is small enough.

We define $p_i = \mathbb{P}[\epsilon = 2^{-i}]$. The approximated unconditional density function for the winning number is then given by:

$$\tilde{f}(x) = \sum_{i=0}^I p_i \sum_{k=0}^{\infty} \tilde{f}_k(x; 2^{-i}) \frac{\tau^k}{k!} e^{-\tau}.$$

Let x_1, \dots, x_M be the sequence of announced numbers in an experiment with M participants.

The log-likelihood function corresponds to:

$$l(\tau, p_1, \dots, p_I) = \sum_{n=1}^M \ln \left(\sum_{i=0}^I p_i \sum_{k=0}^{\infty} \tilde{f}_k(x_n; 2^{-i}) \frac{\tau^k}{k!} e^{-\tau} \right),$$

where $\sum_{i=0}^I p_i = 1$.

$\alpha = 1/2$							
ϵ	y_0^*	y_1^*	y_2^*	y_3^*	y_4^*	y_5^*	y_∞^*
0	25.0	12.5	6.2	3.1	1.6	0.8	0.0
0.25	25.0	14.1	9.2	7.2	6.3	5.9	5.6
0.5	25.0	15.6	12.1	10.8	10.3	10.1	10.0
0.75	25.0	17.2	14.7	14.0	13.7	13.7	13.6
1	25.0	18.7	17.2	16.8	16.7	16.7	16.7

$\alpha = 2/3$							
ϵ	y_0^*	y_1^*	y_2^*	y_3^*	y_4^*	y_5^*	y_∞^*
0	33.3	22.2	14.8	9.9	6.6	4.4	0.0
0.25	33.3	23.6	17.9	14.6	12.7	11.6	10.0
0.5	33.3	25.0	20.8	18.7	17.7	17.2	16.7
0.75	33.3	26.4	23.5	22.3	21.8	21.6	21.4
1	33.3	27.8	25.9	25.3	25.1	25.0	25.0

Table 1: Expected winning numbers y_k^* for several values of α and ϵ . The asymptotic expected winning number y_∞^* is also reported in the last column of the table. The number of players is $M = 1000$.

Model	τ	I	Probabilities							
			p_0	p_1	p_2	p_3	p_4	p_5	p_6	p_7
ϵ uniform	3.02	7.0	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125
$\epsilon \sim (p_i)_{i=0,\dots,I}$	2.99	7.0	0.298	0.123	0.096	0.101	0.098	0.073	0.000	0.211

Table 2: Estimated parameters on the results of the newspaper experiment on the guessing game conducted by Bosch-Domènech and Nagel (1997), which are reported in Figure 1. The first model (ϵ unif) assumes that ϵ is uniformly distributed on $\mathcal{E} = \{2^{-i} : i = 0, \dots, I\}$ and we estimated the parameters τ and I . The second model ($\epsilon \sim (p_i)_{i=0,\dots,I}$) assumes that ϵ is distributed on $\mathcal{E} = \{2^{-i} : i = 0, \dots, 7\}$ and we estimated the parameter τ and the probability distribution $(p_i)_{i=0,\dots,I}$.

session	N4	N5	N6	N7	W	B	N3
α	2/3	2/3	2/3	2/3	2/3	2/3	1/2
ϵ	.99	.42	.76	.71	.35	.2	.29

Table 3: Estimation of the confidence parameter ϵ for several experimental results. N4-7 are Nagel's sessions 4-7 from Nagel (1995) in which $\alpha = 2/3$, W is from Weber (<http://www.andrew.cmu.edu/user/rweber/>), B is from an experiment conducted at the University of Bergen, and N3 is Nagel's session 3 with $\alpha = 1/2$ also taken from Nagel (1995). In Nagel's sessions 1 and 2, the winning number at round 4 of the iterative game is surprisingly much higher than the winning number at round 3 and the log-linear function does not fit the data well.

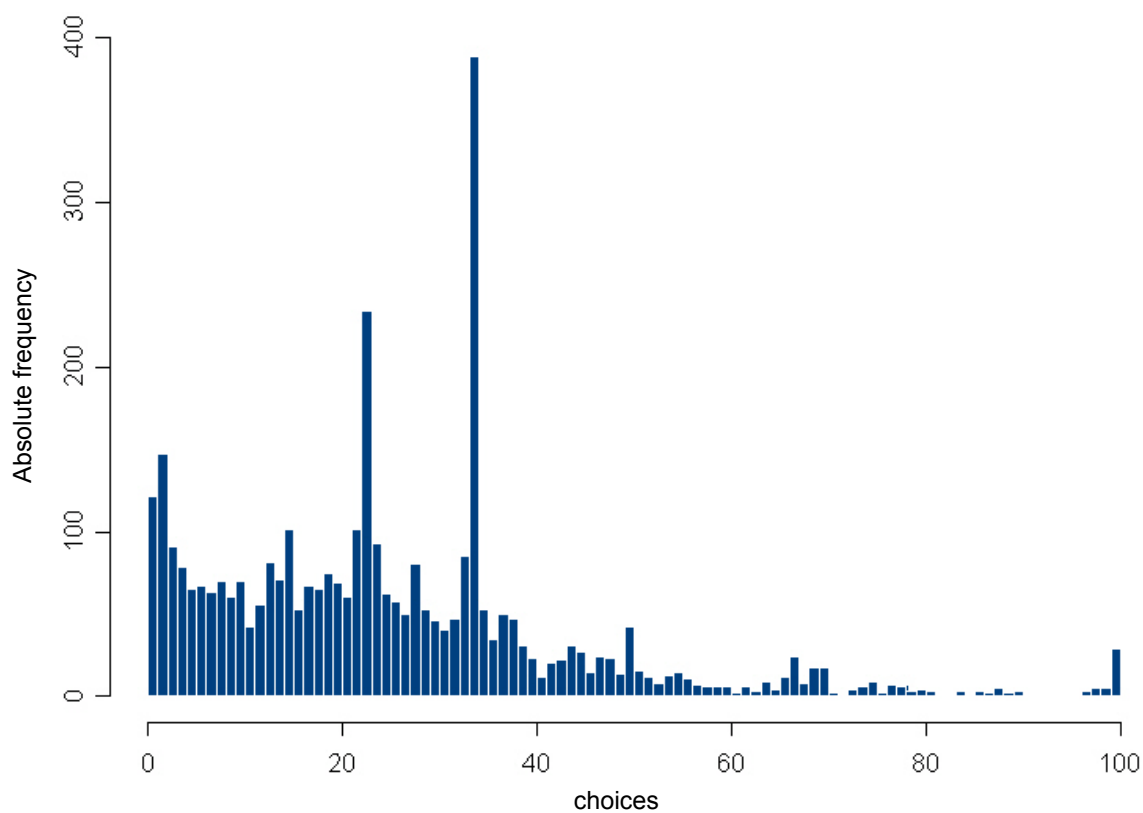


Figure 1: The figure shows the histogram of the results from the one-shot guessing game with $\alpha = 2/3$ played by the Spanish newspaper *Expansion* (Bosch-Domènech and Nagel 1997): 3696 subjects participated in the game. The data set corresponds to the one studied by Nagel, Bosch-Domènech, Satorra, and Garca-Montalvo (2002) and has been provided by Rosemarie Nagel.

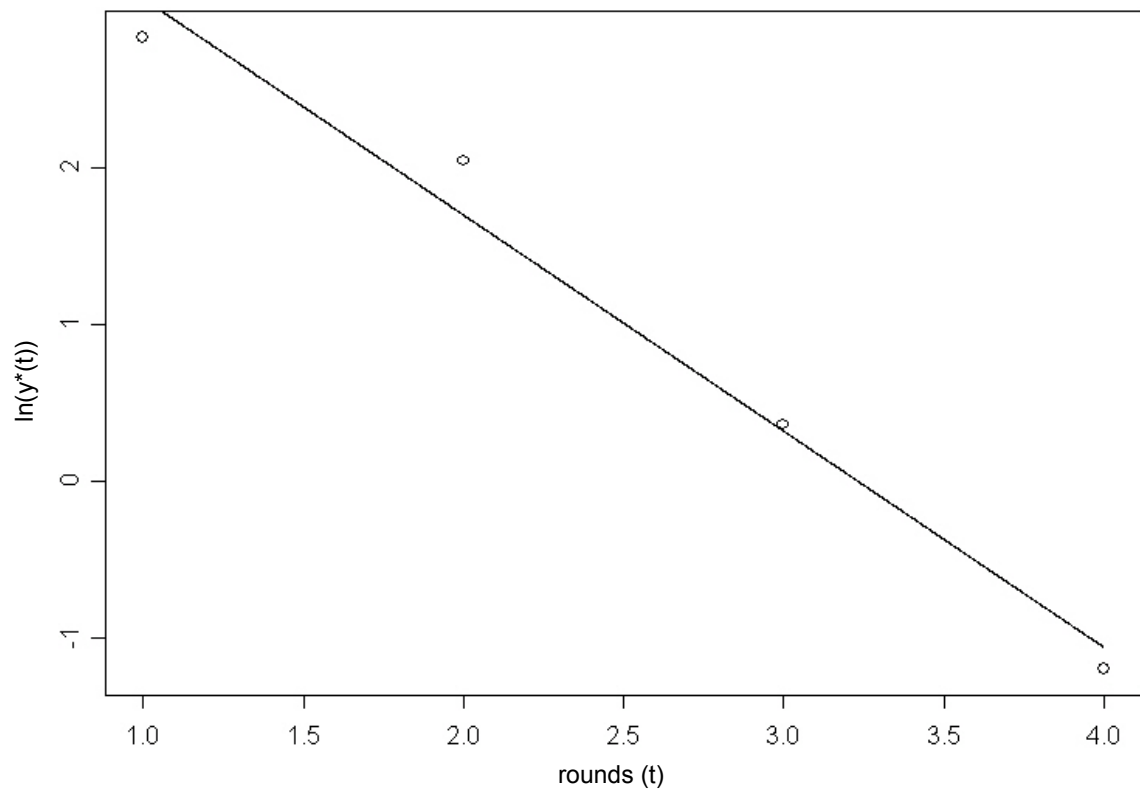


Figure 2: The figure shows the sequence of winning numbers (circles) as function of the number of rounds in the iterated guessing game. The data are from an experiment conducted at the University of Bergen, Norway, 2003. A log-linear function (line) is fitted on the data using OLS estimation ($\ln(y^*(t)) = 4.45 - 1.377t$).

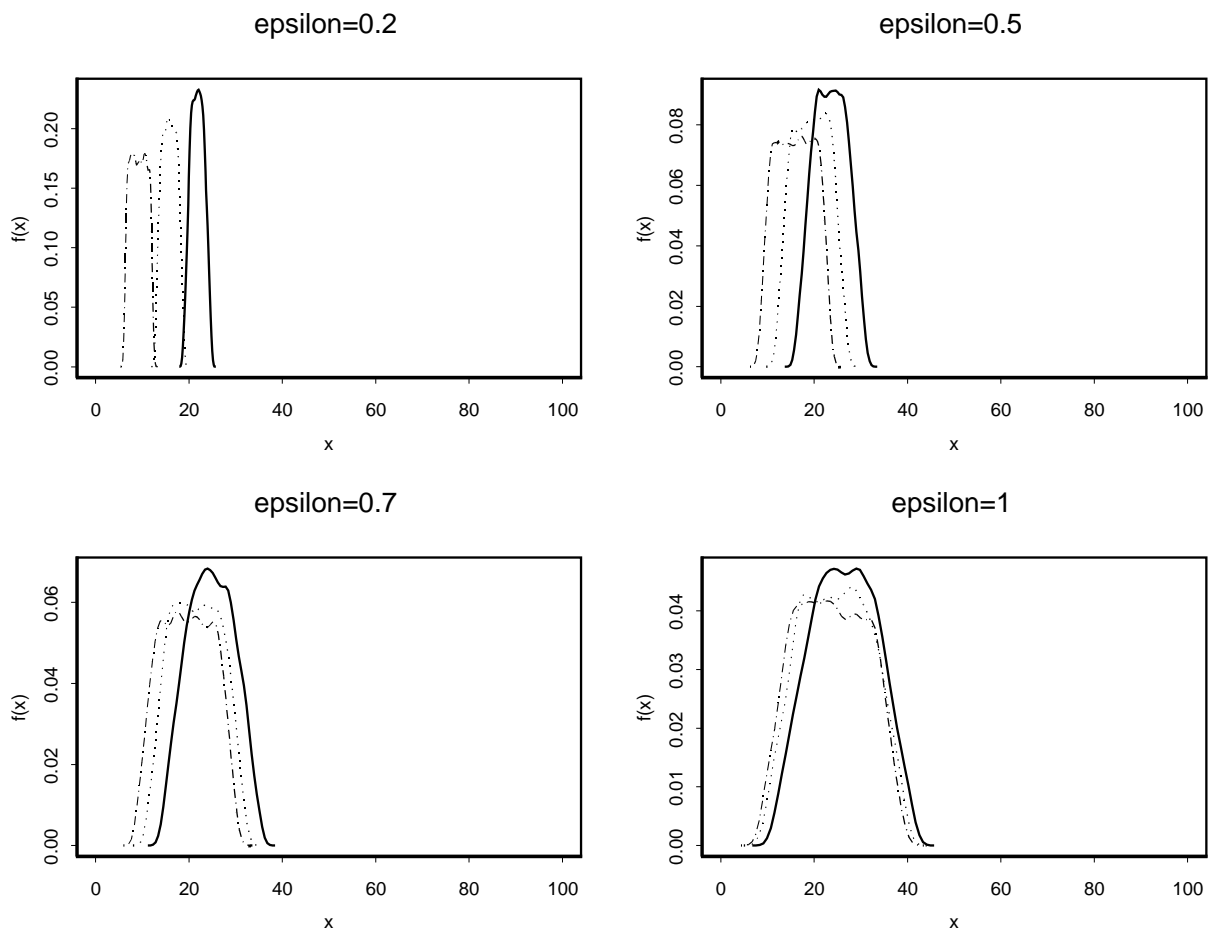


Figure 3: Density function f_k for $k = 1$ (full line), $k = 2$ (dotted line), $k = 5$ (dashed line) and for $\epsilon = 0.2, 0.5, 0.7$ and 1 . The parameters are $\alpha = 2/3$ and $M = 1000$.

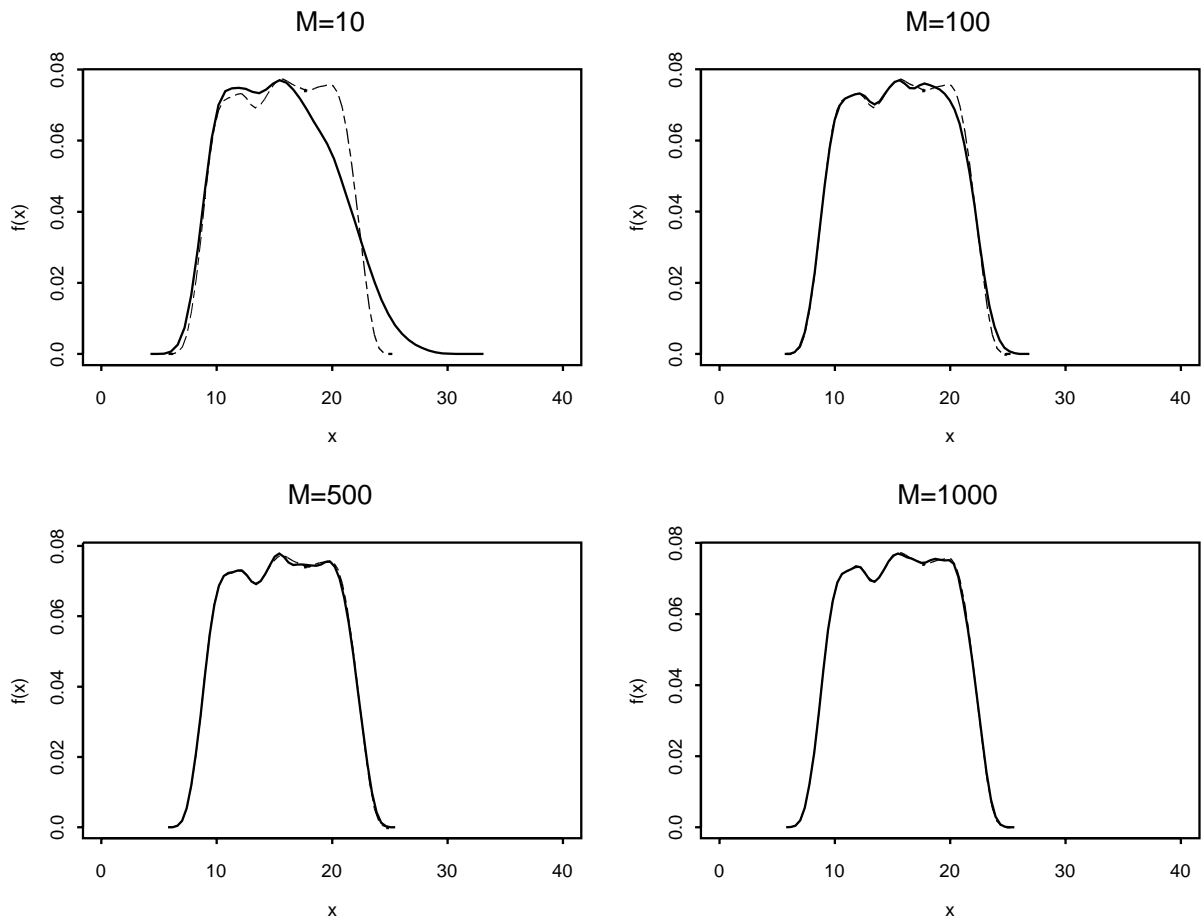


Figure 4: Asymptotic density function f_∞ for $M = 10, 100, 500, 1000$ (full line). The parameters are $\alpha = 2/3$ and $\epsilon = 0.5$. The approximated density function (dashed line, see Appendix B) is also given for comparison's sake. The two functions almost coincide for $M = 500$ and 1000 .

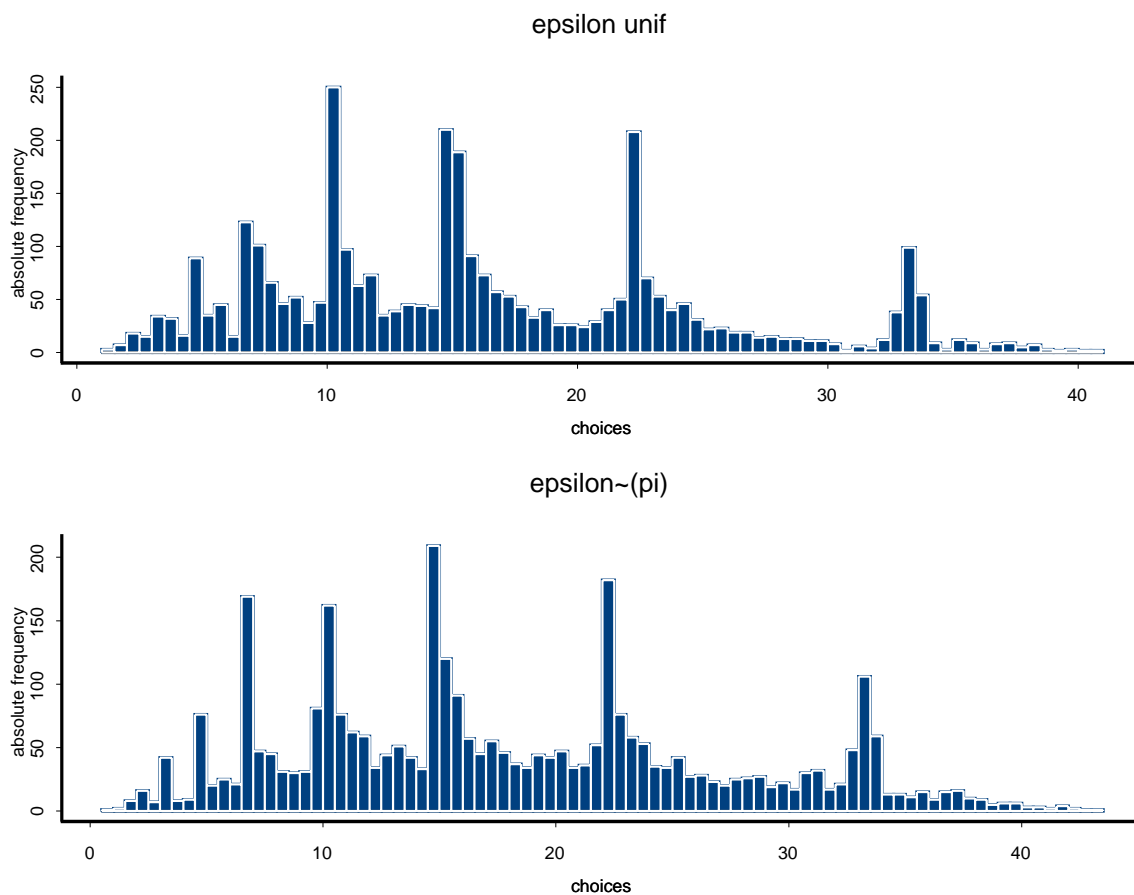


Figure 5: Absolute frequency of announced numbers according to the estimated models on the results of the newspaper experiment on the guessing game conducted by Bosch-Domènech and Nagel (1997), which are reported in Figure 1. The top histogram reports the distribution of announced numbers assuming that ϵ is uniformly distributed on $\mathcal{E} = \{2^{-i} : i = 0, \dots, 7\}$ and $\tau = 3.02$. The bottom histogram reports the distribution of announced numbers assuming that ϵ is distributed on $\mathcal{E} = \{2^{-i} : i = 0, \dots, 7\}$ according to the estimated distribution $(p_i)_{i=0, \dots, 7}$ and $\tau = 2.99$.