

Measuring Risk Attitudes in a Natural Experiment: Data from the Television Game Show LINGO

Citation for published version (APA):

Schotman, P. C., & Beetsma, R. (2001). Measuring Risk Attitudes in a Natural Experiment: Data from the Television Game Show LINGO. *The Economic Journal*, 111, 821-848. <https://doi.org/10.1111/1468-0297.00661>

Document status and date:

Published: 01/01/2001

DOI:

[10.1111/1468-0297.00661](https://doi.org/10.1111/1468-0297.00661)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

MEASURING RISK ATTITUDES IN A NATURAL EXPERIMENT: DATA FROM THE TELEVISION GAME SHOW LINGO*

Roel M. W. J. Beetsma and Peter C. Schotman

We use data from a television game show involving elementary lotteries as a natural experiment to measure risk attitudes. Unique features of our data set are the substantial monetary stakes and the large sample size. CRRA and CARA utility specifications perform approximately equally well. We find robust evidence of substantial risk aversion. Extensions of the basic model, which allow for a separate utility flow purely from playing the game or for decisions based on decision weights instead of actual probabilities, raise the estimated degree of risk aversion.

Risk attitudes play a critical role in determining the demand for insurance as well as the demand for risky assets and the equilibrium returns on these assets. Despite the importance of risk preferences, we still have very little idea about the answer to empirical issues like the average risk aversion of individuals, the determinants of risk attitudes and the heterogeneity of risk preferences among individuals. The empirical findings depend to a large extent on the particular methods used to investigate these issues. Friend and Blume (1975) obtain indirect evidence about the degree of risk aversion from individual asset holdings. They estimate the coefficient of relative risk aversion (RRA) at roughly between 2 and 3. Other indirect evidence is provided by the equity premium puzzle. The observed spread between average stock returns and the risk-free interest rate can only be consistent with an expected-utility maximising framework if we assume extremely large values for the RRA coefficient, usually around 20.¹ Campbell (1996) shows that the puzzle is compounded if one includes human capital in measured wealth.

Direct evidence on the degree of risk aversion is obtained from experiments and surveys. Gertner (1993) uses data from the television game show ‘Card Sharks’, and finds risk-aversion estimates that are comparable to Friend and Blume’s (1975) estimates, although somewhat higher. In another natural experiment, based on the television game show ‘Jeopardy’, Metrick (1995) concludes that players display near risk neutrality. In an experiment in which Chinese students are presented with basic lotteries, Kachelmeier and Shehata (1992) find large differences between how much individuals are willing to pay for a lottery and for how much they are willing to sell the same lottery. This suggests that revealed risk preferences depend on the way problems are

* We thank ID^{TV} and VARA for providing the data. We also thank two anonymous referees, Arjan van Bussel, Eric van Damme, Frank de Jong, Peter Wakker and seminar participants at Stanford University, CentER (Tilburg University), Erasmus University (Rotterdam) and the University of Aarhus for helpful comments. We thank Erik Kole for skillful research assistance. The usual disclaimer applies.

¹ Mehra and Prescott (1985) contains the original statement of the equity premium puzzle. Kocherlakota (1996) and Campbell *et al.* (1997) provide an overview of the literature. The latter also summarise the empirical results for the USA.

framed. Barsky *et al.* (1997) provide survey evidence which suggests that heterogeneity in risk preferences is important.²

In this paper, we use data from the finals of the Dutch television game show LINGO to measure risk attitudes. The LINGO final is a word guessing game which evolves over several rounds. At the start of each round, players decide whether to stop or to continue with the possibility of doubling the money they have won so far or losing everything, in which case the game ends. Due to its specific structure, the game reduces to a series of elementary lotteries in which the survival chances fall over the rounds, while the money stakes rise. The players' decisions thus give an indication about their risk attitudes. For several reasons, LINGO provides a suitable natural experiment in this regard. First, the decisions that players have to take are extremely simple, because LINGO is a non-strategic game. Second, the decisions involve serious amounts of money for most players. On average, a player takes home over 2,500 Dutch guilders (1 Dutch guilder (*f*) is approximately \$0.5), while at the later stages of the game players often must make decisions with more than *f*10,000 at stake. The necessity of large stakes for inference on risk aversion has often been emphasised, most recently by Rabin (2000). Finally, the game has been shown so often on television (on one of the main channels), that most of the players can be expected to be at least reasonably familiar with the game (in addition, there is a pre-qualifying tournament). All this helps to contribute to the validity of the assumption that players take decisions that reveal their true risk preferences.

We start the empirical analysis by investigating to what extent the players' decisions can be explained by 'rules-of-thumb' variables. These are observable statistics from the game (for example, statistics that give an indication of the survival chances, which are not directly observable). By estimating such exploratory regressions, we obtain an idea of what factors are critical in determining the players' decisions.

Our next step is to adopt an expected-utility maximising framework and estimate both constant relative risk aversion (CRRA) and constant absolute risk aversion (CARA) utility specifications for the players. This part of the analysis is in the spirit of Gertner (1993). We confirm his results, qualitatively speaking, with an independent (and larger) data set for a game which is even simpler from the players' perspective.³ If we calibrate initial wealth at the median Dutch household wealth, we obtain an estimate for relative risk aversion of approximately 7.

We consider several extensions of the basic expected-utility model. First, we allow for heterogeneity in the ability to guess words and, therefore, differences in the survival chances across players. This extension hardly affects our results. Second, because LINGO might attract people that are especially keen on

² Heterogeneity in risk preferences may help to explain why so few people are stockholders (Mankiw and Zeldes, 1991), which, in its turn, might have important effects for the behaviour of asset returns.

³ In contrast to Gertner (1993), in modelling the players' decisions, we can take explicit account of the effect of future decisions on the current optimal decision.

playing these types of games and on being on television, we allow for the possibility that the players derive additional utility purely from playing the game. *Ceteris paribus*, teams would continue longer. Hence, to explain the stop decisions in the data, the estimated degree of risk aversion has to increase. Third, we replace the actual survival probabilities with unobserved decision weights, as Prospect Theory – see Kahneman and Tversky (1979; 1982) and Tversky and Kahneman (1992) – suggests. This extension is motivated by the substantial amount of evidence that suggests that individuals tend to be overconfident when taking decisions under risk.⁴ In the current context, individuals may systematically overestimate their odds because they overestimate their ability to solve the word puzzles in LINGO.⁵ Our estimates indeed suggest that players tend to overestimate their survival chances. To nevertheless explain that they do sometimes stop, the estimated degree of risk aversion increases further.⁶

The final step in our analysis consists of some general robustness checks on the model. First, we test whether adding our ‘rules-of-thumb’ variables could help to explain decisions not captured by the expected-utility model. While the latter renders most of these variables irrelevant, it seems that it is not able to account fully for the effects of variables that point to possible heterogeneity in ability and in risk attitude. Second, we simulate the estimated decision model to show that it generates outcomes that match other features of the data.

The remainder of the paper is structured as follows. Section 1 describes the game, while Section 2 describes the data and reports summary statistics. In Section 3 we estimate decision models based on rules-of-thumb. Section 4 provides the theory for the estimation of the expected-utility maximisation models. In particular, we show that due to the particular structure of the game, the players’ multi-stage decision problem can be effectively reduced to a one-stage decision problem. Section 5 reports and discusses the empirical results for the basic expected-utility model. Section 6 explores the extensions of the basic expected-utility model and provides some general robustness checks. Finally, Section 7 concludes the paper.

1. The Game

LINGO is a word game which is shown almost every weekday on Dutch television. By June 1998, it has been played more than 1,600 times. Each show

⁴ Tversky and Kahneman (1974) and Arrow (1982) provide overviews of common biases in people’s assessments of probabilities. Camerer (1987) explores such biases in the context of an experimental market. Gneezy (1997) presents evidence on the failure of individuals to adjust the perceived probability distribution of final outcomes when a game or lottery is played repeatedly.

⁵ Another possibility is that players display bounded rationality – for a recent survey, see Conlisk (1996) – when computing their odds. However, it is hard to see how this could lead to the *systematic* bias in the perceived odds suggested by our results.

⁶ Although it is sometimes argued that transformations of probabilities should be considered as an integral part of an individual’s risk attitudes (Wakker, 1994), we believe that a person’s ability to assess his odds should be distinguished from his personal risk attitude.

involves a 'pre-final' and a 'final'. In the pre-final, two couples have to guess five-letter words (for which they are given certain clues – see Appendix). For each word they guess correctly, a couple receives £50. The couple which has solved the most five-letter words at the end of the pre-final progresses to the final. The other couple goes home with the money it has won and cannot come back for any further games. In the pre-final, teams have one other possibility of winning money, which is totally unrelated to the game. This side effect is called the jackpot (see Appendix).

Our main interest is in the final. The final is played over a maximum number of five rounds. In essence, each round involves an elementary lottery, as depicted in Fig. 1. The couple enters the round with a particular stake. For the first round, this is the amount of money won in the pre-final. At the start of each round, the couple decides whether to continue or to stop. If they stop, they take home their current stake and are not allowed to come back for any further games. If they continue, they must play a lottery. If they win the lottery, their stake for the start of the next round doubles; if they lose, all the money is lost, the game stops and the couple is allowed to come back in the next show unless this is their third, and last, final. If they come back they play again a pre-final against another couple, in which they either drop off, or win and enter the next final.⁷ Given that they are not allowed to come back after their third final, a couple also plays at most three pre-finals. Finally, if a couple has won the lottery in all five rounds in any given final, they are not allowed to come back. Fig. 2 summarises all the possible routes for an arbitrary couple from the start of its first pre-final until it leaves for good. It also shows the *potential* sources of the total accumulated gains for each of the possible routes.

The actual game in each round of the final again involves guessing a five-letter word (with similar clues being given as in the pre-final). The maximum number of attempts for a word in a given round is five. The number of attempts determines the number of balls a couple has to draw from an urn. In

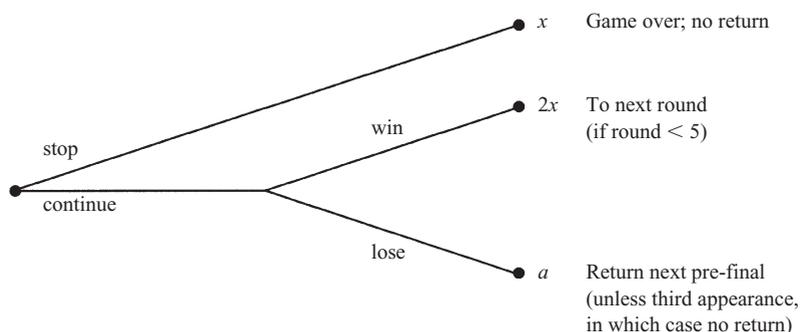
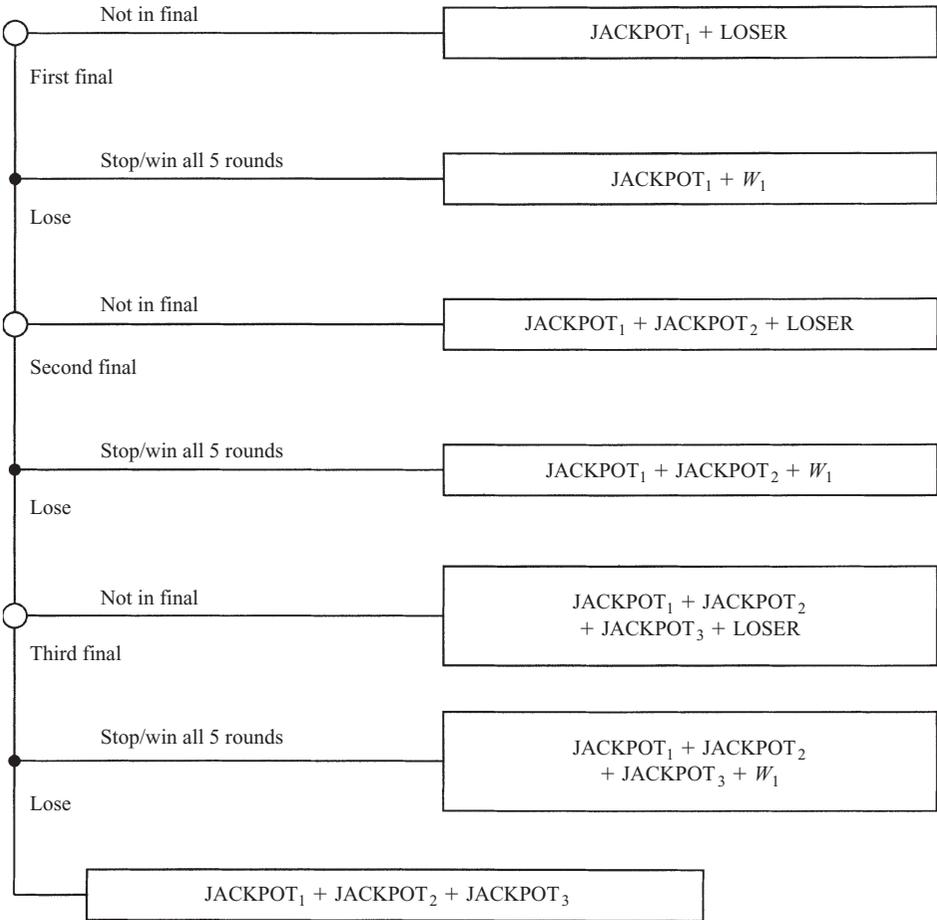


Fig. 1. *Lotteries and Decisions*

⁷ The next show is usually recorded within an hour's time, but is shown on television the next day.



The figure depicts all the possible routes for an arbitrary couple from the start of its first pre-final until it leaves LINGO for good (which happens upon reaching one of the endpoints in the chart). It also shows the *potential* sources of gains for each of the possible routes. LOSER denotes the earnings of a couple that drops out in the pre-final. W_1 is the amount won in the final. JACKPOT₁, JACKPOT₂, JACKPOT₃ are the jackpot gains in the first, second and third pre-final, respectively. Open circles denote the prefinals.

Fig. 2. Flow Chart and Possible Sources of Gains

particular, for each attempt needed to guess the word in some round the couple has to draw one ball (without replacement) from the urn. If they did not guess the word in five attempts, they must draw the maximum number of six balls.

Now consider the LINGO board in Fig. 3. At the start of the first round in the final, each of the nine *open* boxes, i.e., those not filled with an \times , contains some number, while the urn contains 35 numbered balls and one *golden* ball. If the number of a ball the couple draws from the urn corresponds to the

X	1	X	2	X
3	X	X	X	4
X	X	5	X	X
6	X	X	X	7
X	8	X	9	X

Fig. 3. *Example of a Lingo Board*

number in one of the open boxes, this box is (and remains) filled with an \times .⁸ Once one row, column or the diagonal of the board is completely filled with \times s, the game is over the players lose their stake, as explained above. In the initial position this would happen if the ball with the number 5 is drawn.

Once the couple has drawn the required number of balls (without a row, column or the diagonal of the board having become completely filled with \times), they go to the start of the next round. If, at some point, they draw the golden ball, they stop drawing balls and survive the round. The golden ball is placed back into the urn and the couple go immediately to the start of the next round. At the start of the next round, the stake is doubled and they make their new decision whether to stop (and take home their new stake) or to continue. The starting position of the board and the urn is the one inherited from the end of the previous round.

Summarising, LINGO provides a natural experiment for measuring risk attitudes, because at the start of each round in a given final, the players face a very simple decision problem: whether to continue. Moreover, because with any new draw from the urn, either the number of 'good' balls in the urn (ie, the balls with a number that is not on the LINGO board) decreases or another box of the board is filled with an \times , the survival probabilities are decreasing over the rounds. This should induce players to stop if the odds become too unfavourable.

⁸ For each of the nine numbers in the open boxes, there is exactly one ball in the urn which contains the same number.

2. Data and Descriptive Analysis

Most of the data were generously provided by the producer, ID^{TV}. The remainder of the data were obtained by watching the programme on television.⁹ We use observations from finals 633 to 1,611. This is a total of 979 finals but, of course, this constitutes a much larger number of decisions, given that each final can last for up to five rounds.

Table 1 shows some key summary statistics for the finals. The entries only refer to the players who appear in the final. Candidates are selected on the basis of their ability to guess words and their appearance on television. Hence, there is no apparent link between the selection criterion and their risk attitude. For most of the couples, we know their gender. Females are clearly overrepresented in this game. For about 20% of the finals, we know the ages of the players. By far, most of them are in their twenties or thirties. Players younger than twenty and older than sixty-five are very rare. We also know the education level and/or profession for about 20% of the participants. From this

Table 1.
Summary Statistics

Variable	Average	Standard deviation	Min.	Max.	No. obs
<i>Statistics per show</i>					
SEX	0.2	0.8	-1	1	961
AGE	32.3	8.7	18	70	401
LOSER	361	643	0	630	965
JACKPOT	2,327	2,819	0	20,100	965
FINAL	1.4	0.6	1	3	979
W_0	473	106	150	950	979
W_1	2,798	4,962	0	27,200	979
WINJACK	447	1,485	0	20,450	967
<i>Statistics per team</i>					
SEX	0.2	0.8	-1	1	489
AGE	32.5	9.1	18	70	271
W_1	4,106	5,547	0	27,200	667
TOTAL	4,973	5,510	0	28,650	667

Notes: *SEX* describes the sex composition of a couple: -1, if both players are male, 1, if both are female; and 0, otherwise. *AGE* is the age of a participant (the number of observations refers to the number of individuals for which we know *AGE*, not the number of teams). *LOSER* is the money won by the losers of the pre-final (including the money they win from the jackpot). *JACKPOT* is the amount of money in the jackpot at the end of the pre-final. *FINAL* is the number of finals played including this one. W_0 is the initial stake in the final; W_1 are the total gains in the final. *WINJACK* are the gains from the jackpot in the pre-final. *TOTAL* gives the total gains of a team, possibly accumulated over different games and including jackpot wins. The first part of the table pertains to all games; the second part gives statistics per couple. The number of observations can differ among variables due to missing values in the early games. All monetary variables are in Dutch guilders.

⁹ ID^{TV} keeps a record of the course of each final. These records are used by its accountant for the official tax statement. These data should be of high quality, which is confirmed by comparing them with our observations from television.

it appears that students are overrepresented. We have little idea how the composition of the sample will affect our measures of risk aversion. One source of information is Hartog *et al.* (2000) who conducted surveys among Dutch people, in which respondents were asked about the amounts they would be willing to pay for certain lottery tickets. From the answers, it appears that women are more risk averse than men and that the schooling level reduces risk aversion. The initial stake in the final, which is equal to the amount a team wins in the pre-final (W_0 guilders), is on average $f473$. This translates to ($2^5 \times f473 =$) $f15,136$ potential gains for a team that plays and wins all five rounds in the final. The average realised total gains per couple is approximately $f5,000$, which would be quite a serious amount of money for most of them. The distribution of the gains is very dispersed; more than half of the teams leave without winning anything in the final.

An important input for calculating the odds in the lotteries is the ability of couples in solving the word puzzles. Table 2 gives the frequency distribution of the number of balls that had to be drawn in each of the rounds of the observed finals. We tested for many different ways of splitting the sample to search for heterogeneity in ability. Gender differences, age or number of appearances in the final do not significantly alter the ability distribution. The null hypothesis of the distributions being equal could only be rejected for a split according to the amount of money a couple has won in the pre-final round (W_0). The table shows the distributions for $W_0 \leq > 400$. Even here the differences are small. Since we cannot find reliable other instruments for possible heterogeneity in ability, unless explicitly stated otherwise, we make the assumption that all couples have the same ability in solving the word puzzles.

Table 3 separates the first, second and third finals into the percentage of couples that continued and survived all five rounds (*WINNER*), those who stopped before or at the end of the fourth round (*STOPPER*), and those who won nothing this final (*LOSER*). As is clear from the table, the percentage of losers falls, while the percentage of stoppers increases with the number of appearances in the final. This indicates that people are less willing to take risks, if they run out of possibilities to return in future shows. Of the losers in the first final 49% ($= 232/475$) managed to come back into the second final,

Table 2
Frequency Distribution of Balls to be Drawn

	1	2	3	4	5	6
OVERALL f_j	0.14	0.32	0.23	0.13	0.081	0.089
$W_0 \leq 400$	0.15	0.29	0.24	0.14	0.081	0.100
$W_0 > 400$	0.14	0.34	0.23	0.13	0.079	0.080

Notes: The entries report the frequency of the number of balls to be drawn from the urn per round. The OVERALL line is the average over all plays. The other lines split the sample according to initial stakes ($W_0 \leq > 400$). The hypothesis that abilities do not depend on W_0 is rejected at the 5% level: $\chi^2(5) = 12.38$ ($p = 0.03$).

Table 3
Winners, Losers and Stoppers in Finals

FINAL	1		2		3	
	Abs	%	Abs	%	Abs	%
WINNER	70	11	18	8	5	6
STOPPER	122	18	55	24	31	39
LOSER	475	71	159	69	44	55
TOTAL	667	100	232	100	80	100

Notes: Entries report the number of finals that end with a couple winning all five rounds (WINNER), couples deciding to stop (STOPPER), and couples ending the final with nothing (LOSER). Abs is the absolute number of finals, and % the percentage of the total.

while 50% (= 80/159) of the losers in the second final returned in the third final. As this means that the probability of winning against a random couple in the next pre-final is about one half, there must be only minor differences in abilities across the couples participating in the game.

An important input for the decisions are the odds at which couples play the next lottery. If j balls must be drawn, then the probability of surviving the next round can be calculated from the number of balls that remain in the urn (BALLS) and the number of direct possibilities of filling a row, column or the diagonal on the LINGO board (BAD). This probability is denoted $P_j(\text{BAD}, \text{BALLS})$. BAD and BALLS together denote the *state* of game. Using the ability distribution f_j from Table 2 we compute the overall probability of surviving the next round as

$$\text{PROB} = \sum_{j=1}^6 f_j P_j(\text{BAD}, \text{BALLS}). \tag{1}$$

Table 4 shows summary statistics of the probabilities at which the players decided to stop, and at which they made their last play decision. On average couples stop at survival probabilities well above a half. We almost never observe them playing at probabilities less than one half. Some couples are apparently so risk averse that they stop at a survival probability of 0.86.

Table 4
Odds at Decision Moments

	Average	Standard deviation	Min.	Max.
STOPPERS	0.62	0.08	0.44	0.86
PLAYERS	0.71	0.11	0.46	0.88

Notes: The entries show the survival probabilities when couples made their last play decision, or when they stopped.

3. Decisions Based on Rules-of-Thumb Variables

This section contains an exploratory data analysis to find out which variables players may use in their decisions. Before turning to decision rules based on expected utility maximisation (in the next section), we will study possible 'rules-of-thumb' variables that players might use in deciding whether to stop ($y_i = 0$) or to continue ($y_i = 1$). We use Probit regressions based on an indicator y_i^* which is a linear function of variables that players can observe immediately or that they can compute with little effort. We consider variables that can reasonably be expected to play a role in the decisions (for example, because they provide an indication of the survival chances). Formally, the players stop if $y_i^* < 0$ and continue otherwise.

The most interesting decisions are the last decision in each show and, in shows where the players decided to stop, also the next-to-last decision. The last decisions of winners and losers give an indication of an upper bound on their risk aversion, while the final two decisions of stoppers jointly give an upper and a lower bound on their risk aversion. For a given final, our data set of decisions thus includes stoppers twice (the round they stop and the round before that, when they still decided to continue) and the other couples once (their last 'play' decision, whether they win or lose). The total number of observations is 1,186. We assume that the error terms in the Probit regressions are uncorrelated, even though we sometimes observe multiple decisions by the same couple.¹⁰ These are the explanatory variables for the decision to play:

- *STAKE*: the amount of money at stake at the moment the decision is to be taken
It is included because, at given odds, a larger stake increases the riskiness of the bet as measured by the variance of the possible outcomes. For this reason, we would expect *STAKE* to have a negative effect on the likelihood to continue in this final.
- *FINAL*: the number of finals (including this one) played by the candidates.
If *FINAL* is higher, the number of possibilities for coming back is reduced, which should discourage the players from continuing. This suggests that *FINAL* has a negative effect.
- *BALLS*: the number of balls still in the urn
A larger value of *BALLS* raises the survival probability and, hence, we expect its effect to be positive.
- *BAD*: the number of open positions on the LINGO board for which a row, column or the diagonal would be directly filled if the corresponding number were drawn from the urn
BAD is expected to have a negative effect.
- *LAST*: the last round of play in this final
This variable is included as an indication of heterogeneity. If teams differ in their risk attitudes, then the more risk-averse players will stop earlier.

¹⁰ The possible correlations will have no easy pattern, since we can have multiple observations of the same team for several reasons. A couple may decide to stop (2 observations) or observations span multiple shows (when a couple returns in a second or third final).

That means that in later rounds of the game, players with less risk aversion are overrepresented. Because they are less risk averse on average, we expect teams to be more willing to play in later rounds, *ceteris paribus*.

- W_0 : the initial stake in this final, which may serve as a proxy for the ability to solve word puzzles

Ceteris paribus, better players would have an incentive to stop later, which suggests that it should have a positive effect.

- *PROB*: the survival probability (computed according to (1))

If players in fact base their decisions on a rule-of-thumb, then *PROB* should not have any independent explanatory power. It might also be too complicated to calculate or estimate *PROB* on the spot. Since *PROB* is a function of *BAD* and *BALLS* (though highly nonlinear), players who could somehow estimate their survival probabilities, would no longer need to base decisions on *BAD* or *BALLS* separately, once they know *PROB*. This suggests that either *PROB*, or the combination of *BAD* and *BALLS* is redundant (as an explanatory variable).

Table 5 contains the results of the Probit regressions. All explanatory variables are significant and have the expected sign, although *PROB*, *BAD* and *BALLS* cannot be included simultaneously due to severe multicollinearity. The first regression in Table 5 does not include *PROB*, and thus tries to explain decisions by employing only variables that are directly observable by the players. In the second regression we replace the regressors *BALLS* and *BAD* by *PROB*. While the other coefficients are basically unaffected, *PROB* becomes highly significant. The fit of this model is hardly different from the fit of the first regression. These results suggest that players take account of their odds by basing their decisions on variables which are highly correlated with the survival probabilities computed with the help of equation (1).

This exploratory model serves as a benchmark to evaluate the more structural expected-utility maximisation models that we develop below.

Table 5
Probit Regressions with Rules-of-thumb Variables

<i>C</i>	<i>PROB</i>	<i>STAKE</i>	<i>FINAL</i>	<i>BALLS</i>	<i>BAD</i>	<i>LAST</i>	W_0	$\ln L$	R^2
-1.02 (1.05)	-	-0.38 (0.04)	-0.39 (0.08)	0.09 (0.02)	-0.17 (0.05)	0.31 (0.11)	2.47 (0.65)	-350.00	0.36
-1.53 (0.66)	3.81 (0.62)	-0.38 (0.04)	-0.39 (0.08)	-	-	0.25 (0.10)	2.60 (0.64)	-352.78	0.36

Notes: Dependent variable is the play/stop decision (1/0). Explanatory variables are a constant (*C*), the probability to survive next round (*PROB*), the prize money at stake (*STAKE*), the number of times the team appeared in the final including this final (*FINAL*), the number of balls still in the urn (*BALLS*), the number of direct possibilities to fill a row, column or diagonal (*BAD*), the round of play in this final (*LAST*) and the amount with which the team enters the final (W_0). Standard errors are in parentheses. In the regressions the monetary variables (*STAKE*, W_0) have been divided by 1,000. The column in ' $\ln L$ ' reports the maximised value of the loglikelihood function, and R^2 is the pseudo- R^2 for probit regressions. The number of observations is $N = 1,186$.

4. Expected-Utility Maximisation

We turn now to the analysis of decision models based on expected-utility maximisation. We are interested in how well decisions can be explained using utility functions that are often employed in economic analysis. Moreover, we would like to know how well these models perform when compared with the decision models based on rules-of-thumb, which we estimated in Section 3. The main objective of the expected-utility analysis is to obtain estimates of the degree of risk aversion.

The decision problem faced by the players seems complicated, because it may be followed by other decisions. One such complicating factor is the potential possibility to return the next show after losing today. Instead of getting zero, one can expect to gain something from tomorrow's game. The return possibility can be seen as an option which affects decisions in this final whether to stop or continue. The other option element is the possibility of deciding again in the next round of the current final. For example, suppose that the game has evolved until the start of the second round. The decision to stop would eliminate the possibility of making a similar decision at the start of the third round and could therefore affect the decision taken at the start of the second round. In general, this aspect would require dynamic programming techniques to solve for the optimal stopping rule.

In this section, we provide some theoretical results pertaining to the specific game under consideration. In particular, we show that, because of the (weakly) decreasing survival probabilities over the rounds, the decision problem reduces, under rather weak conditions, from a multi-stage dynamic programming problem to the choice between receiving the current stake with certainty and a single elementary lottery. This simplifies the analysis of the game considerably, and motivates the econometric model.

We start the analysis by considering a one-round game, which could be the last round of an actual final. Let $U(x)$ be the utility that a couple gets from receiving amount x .¹¹ Fig. 1 (page 824) shows the decision tree for the one-round game. At the start of the round the stake is x . The probability of surviving this round (PROB) is denoted p . If the players decide to stop, they keep x with certainty. If they continue, they receive $2x$ with probability p , while they obtain utility $U(a)$ with probability $1 - p$, where $U(a)$ depends on how often they could possibly play the final in future shows. The amount a therefore represents the certainty equivalent (expressed in guilders) of coming back. In other words, it is the amount they would be willing to pay to come back in the next show. If the players cannot come back (because it is their third final), then $a = 0$.

Denoted by $V(A)$ the expected utility from action $A = 0$ (stop) or $A = 1$ (continue) at the start of a hypothetical one-round game, so that

¹¹ This assumes that the two players that form a couple have the same utility function, which implies that they do not disagree about the decision to be taken. In fact, when watching the show on television we very rarely observe disagreement about whether to continue or to stop. Moreover, we assume that they share their gains equally. Because we have no way of knowing how they actually share their gains, this seems to be a rather natural assumption.

$$V(0) = U(x)$$

$$V(1) = (1 - p)U(a) + pU(2x) \quad (2)$$

If $x \leq a$, it is always optimal to continue. Throughout this section, we will therefore assume that $x > a$. The players will decide to stop if $V(1) < V(0)$, or equivalently if p is smaller than the critical probability p^* given by

$$p^* = \frac{U(x) - U(a)}{U(2x) - U(a)} \quad (3)$$

Observe that p^* is decreasing in $U(a)$, which means that continuing becomes relatively more attractive the higher is the utility from returning the next day. If $x < a$, we define $p^* = 0$, and players will always continue.

In the empirical analysis we will use two different utility functions. The first is the constant relative risk aversion (CRRA) specification

$$U(x) = \frac{(W + x)^{1-\gamma} - 1}{1 - \gamma} \quad \gamma > 0 \quad (4)$$

where W acts as a wealth level against which the gains (x) in this final will be evaluated. We assume that $W \geq 0$ and $\gamma > 0$. Moreover, if $W = 0$, then we impose $0 < \gamma < 1$ (otherwise $U(0)$ would no longer be defined), while $\lim_{\gamma \rightarrow 1} U(x) = \ln(W + x)$. The critical probability p^* follows as

$$p^* = \frac{(W + x)^{1-\gamma} - (W + a)^{1-\gamma}}{(W + 2x)^{1-\gamma} - (W + a)^{1-\gamma}} \quad (5)$$

showing that the optimal stopping rule will depend on all parameters (x, W, γ, a).¹²

W can have several possible interpretations and will, therefore, be referred to as the *reference wealth level*. Strictly speaking, under expected-utility maximisation W would be interpreted as the wealth level at the start of the final. Because initial wealth may well differ across couples and because it is unobservable, it is less suitable in the empirical work. When we use the CRRA function, we must therefore treat W as a parameter and interpret it as the average level of wealth of our sample. This is an important drawback of the CRRA utility function.

In the prospect theory developed by Kahneman and Tversky (1979), individuals attach utility, or more precisely 'value', only to current gains and losses rather than to final wealth. In that case, $W = 0$ and the CRRA specification reduces to the power utility function

$$U(x) = \frac{x^{1-\gamma} - 1}{1 - \gamma} \quad 0 < \gamma < 1 \quad (6)$$

In their experiments, Kahneman and Tversky (1982) find support for equation

¹² If a couple shares their gains equally, each would have a utility $U(x/2)$, and an initial wealth $W/2$. Since (5) is homogeneous of degree zero in x, a and W , decisions based on individual utility or team utility will be the same.

(6) as a description of individuals' evaluation of gains ($x > 0$). A very special case is the power utility function for teams in the third final (when $a = 0$). In that case, $p^* = (\frac{1}{2})^{1-\gamma}$, so that decisions would be independent of the stakes of the lottery. In all other cases, the stakes will matter.

The other type of utility function that we consider is constant absolute risk aversion (CARA):

$$U(x) = -\frac{1}{\gamma} \exp[-\gamma(W + x)] \quad (7)$$

Parameter γ now denotes the coefficient of absolute risk aversion. Using the data from their survey, Hartog *et al.* (2000) conclude that CARA is a reasonable description of the risk attitude of individuals with financial wealth of up to f100,000. This is substantially higher than the median household wealth in the Netherlands.

The different implications of the CARA utility function can be seen by computing the critical probability

$$p^* = \frac{\exp(-\gamma x) - \exp(-\gamma a)}{\exp(-2\gamma x) - \exp(-\gamma a)} \quad (8)$$

which does not depend on wealth W . Risk neutrality corresponds to the limiting case with $\gamma \downarrow 0$. To see this, apply l'Hôpital's rule to (8), to obtain,

$$\lim_{\gamma \downarrow 0} p^* = \frac{x - a}{2x - a}. \quad (9)$$

The right-hand side of this equation is the probability p for which a risk-neutral person is indifferent between receiving x with certainty, or $2x$ with probability p and a with probability $(1 - p)$.

Next consider Fig. 4, which shows a two-round game. It could represent the last two rounds within a given final. Hence, the option value a remains constant over the two rounds. At the start of the first round, the couple either stops and takes home x with certainty, or they continue. In the latter case, they survive, with probability p_1 and lose with probability $1 - p_1$. If they play and survive, they enter the second round, where the game is of the same format as the one-round game analysed earlier in the section, although now with a probability p_2 and an initial stake of $2x$. The expected value of coming back, given by $U(a)$, is the same as in the first round. The survival probability p_2 at the start of the second round is unknown at the start of the first round. Hence, the expected utilities from stopping or continuing at the start of the first round of the two-round game are given by

$$V_2(0) = U(x)$$

$$V_2(1) = (1 - p_1)U(a) + p_1E\{\max[U(2x), (1 - p_2)U(a) + p_2U(4x)]\} \quad (10)$$

where the expectation in (10) is taken over the unknown survival probability p_2 at the start of the second round. The subscript on the value function $V_n(\cdot)$

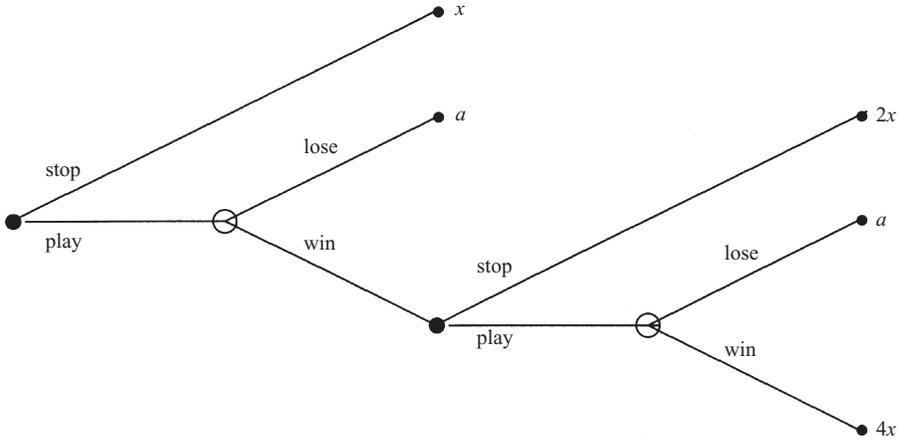


Fig. 4. *Two-round Game*

indicates that we are evaluating at the start of an n -round game. Hence, $V_1(A) = V(A)$ in (2). It follows immediately from (10) that

$$\begin{aligned}
 V_2(0) &= V_1(0) = U(x) \\
 V_2(1) &\geq (1 - p_1)U(a) + p_1U(2x) = V_1(1)
 \end{aligned}
 \tag{11}$$

Hence, if it is optimal to continue at the start of the one-round game ($V_1(1) > V_1(0)$), it is also optimal to continue at the start of the two-round game.

Define the round- k critical probability p_k^* as

$$p_k^* = \frac{U(2^{k-1}x) - U(a)}{U(2^kx) - U(a)} \quad k = 1, \dots, n
 \tag{12}$$

The critical probabilities have the following useful property:

LEMMA 1. *For both CRRA and CARA utility functions, the critical probabilities weakly increase per round:*

$$p_{k+1}^* \geq p_k^* \quad k \geq 1
 \tag{13}$$

Also, for the CRRA function, for given x , p^* is decreasing in W .

Proof of the lemma follows by straightforward algebra, and is available on request. The lemma does not hold in general for any utility function, though. We can now state the main result of this section, which is summarised in the following proposition:

PROPOSITION 1. *Suppose (13) holds. The optimal decision at the start of the first round of an n -round game is to continue if and only if $p_1 \geq p_1^*$.*

Proof Here, we give the proof for the two-round game only. This serves as the first step of the induction argument for the general case, which is available upon request. We have already seen in (11) that, if it is optimal to continue at the start of the one-round game, it is also optimal to continue at the start of the two-round game. Now, we provide the proof for the stop decision. Suppose that $p_1 < p_1^*$, so that it would be optimal to stop at the start of the one-round game. Since $p_2 \leq p_1$ and $p_1^* \leq p_2^*$ (by assumption), one has that $p_2 \leq p_2^*$. Hence, by the definition of p_2^* as the critical probability, we have

$$\max[U(2x), (1 - p_2)U(a) + p_2U(4x)] = U(2x)$$

for all possible realisations of p_2 . Therefore,

$$V_2(1) = (1 - p_1)U(a) + p_1U(2x) = V_1(1) < V_1(0) = V_2(0)$$

Together, Proposition 1 and Lemma 1 imply that, for CRRA and CARA utility, a team should stop the first time the elementary lottery becomes unfavourable. The option to make similar decisions in future rounds of a given final is worthless. The intuition is that, if the odds are too unfavourable to continue now, they will certainly be too unfavourable in future rounds, because they cannot improve (the set of open positions on the LINGO board remains constant or shrinks in the remainder of a given final) and the critical survival probabilities weakly increase (because the amount of money at stake rises).

5. Empirical Results for the Expected-Utility Model

In this section, we present estimates of decision models based on expected utility maximisation. As we already noted, we assume that all players have the same utility function. For now, we assume also that all couples have equal abilities in guessing five-letter words and, therefore, equal survival probabilities (for a given state of the game).

In Section 4 we showed that, due to the special structure of the game, under very weak restrictions (see Proposition 1), the player's decision problem at any moment reduces to an elementary lottery. Expected-utility maximisation then motivates the following probit model for the stop/play decision,

$$y_i^* = p_i - p_i^* + \epsilon_i \quad (14)$$

where p_i is the survival probability computed as the right-hand side of (1), and with observations $y_i = 0$ (stop) if $y_i^* < 0$ and $y_i = 1$ (continue) if $y_i^* \geq 0$. Teams will play if the survival probability p_i exceeds the critical survival probability p_i^* . The error term ϵ_i is assumed to be normally distributed with mean zero and variance σ^2 . The parameters of the model are the degree of

risk aversion γ , the option value of returning the next show for teams in the first or second final (a_1 and a_2 , respectively), reference wealth W (for CRRA utility), and the error variance $\sigma^2 = E(\epsilon_i^2)$.^{13,14} The error term could capture a variety of effects. One would be the presence of non-systematic errors in the assessment of the survival probability p_i . Another would be that individuals make mistakes in their choices even when they know their odds.¹⁵ Finally, the error term may arise from the presence of heterogeneity among the players which is not systematically linked to their survived probabilities.

Table 6 shows the estimates for the CRRA utility specification. The first row shows the results for a power utility specification ($W = 0$). This specification assumes that players' utility is determined only by the gains from the game, without any reference to their level of wealth. The estimate of γ of 0.42 is highly significant (conditional on $W = 0$), and indicates a strong rejection of risk neutrality ($\gamma = 0$).

Table 6
Constant Relative Risk Aversion

	Wealth	γ	a_1	a_2	σ	N	$\ln L$
All games	0	0.42 (0.05)	1,801 (219)	1,493 (229)	0.24	1,186	-380.77
	2,000	0.97 (0.09)	1,882 (183)	1,565 (206)	0.20	1,186	-359.76
	10,000	2.06 (0.20)	1,896 (194)	1,552 (221)	0.21	1,186	-353.82
	20,000	3.31 (0.33)	1,893 (199)	1,542 (236)	0.20	1,186	-352.65
	50,000	6.99 (0.72)	1,890 (204)	1,532 (232)	0.21	1,186	-351.92
	100,000	13.08 (1.35)	1,888 (206)	1,528 (233)	0.21	1,186	-351.69
Third finals	50,000	6.03 (1.33)			0.16	111	-50.35
NOT third finals	50,000	7.55 (0.91)	2,077 (225)	1,691 (256)	0.23	1,075	-299.63

Notes: Dependent variable is the play/stop decision (1/0). γ is the coefficient of relative risk aversion in the CRRA model. 'Wealth' denotes the reference level of wealth. a_1 and a_2 are the equivalent (guilder) amounts of the 'comeback' utilities. $\ln L$ is the value of the loglikelihood at the optimum, and σ is the standard deviation of the error term. Standard errors are in parentheses. The parameters have been estimated by maximum likelihood conditional on different values of reference wealth. The lower panel of the table is based on the subsample of teams that are in the final for the third (and last) time, and the complementary subsample. The hypothesis of equality of γ in the two subsamples cannot be rejected at the 5% level.

¹³ While in the previous section, we kept a constant over the rounds, because we considered a given final, here a is allowed to vary according to whether observations belong to first or second finals (for third finals, $a = 0$). We expect that $a_1 > a_2$, because the possibility of coming back after the first final also allows for a possibility of coming back after the second final.

¹⁴ Contrary to the standard probit model the residual standard error σ is identified. Equivalently, we could have written $y_i^* = \beta(p_i - p_i^*) + \epsilon_i$ with the variance of ϵ_i normalised to one.

¹⁵ In particular, we may observe a stop decision even when $x \leq a$. Because stop decisions at low stakes are very rare, this happens only in a handful of cases.

The other rows of the table report estimates of γ for different levels of reference wealth to demonstrate the dependence between γ and W . When we increase the wealth parameter, the estimate for γ also increases. With wealth equal to $f50,000$, the relative risk aversion parameter increases to 6.99. All these models with wealth in excess of $f10,000$ fit the data about equally well. This multicollinearity arises because the informative stakes in the games are mostly concentrated in a range between 2,000 and 5,000 guilders. Within this range, the utility functions $-(10,000 + x)^{-1.06}$ and $-(50,000 + x)^{-5.99}$ imply very similar critical probabilities p_i^* .

This is illustrated by the following example. Consider a lottery with $f4,000$ at stake ($x = 4,000$). A person with a low wealth $W = 10,000$ will play the lottery if the probability of winning is larger than $p^* = 0.65$. Someone with the higher wealth $W = 50,000$, but the same relative risk-aversion coefficient will always play the lottery at this probability, since the stake is relatively low compared to his higher wealth. To obtain a critical probability $p^* = 0.65$ when the reference wealth level is $f50,000$, a person must have a higher risk-aversion coefficient. Indeed, when $\gamma = 6.99$, the critical probability for the $f4,000$ lottery is 0.63, almost equal to the critical probability at ($W = 10,000$, $\gamma = 2.06$). Over the range of stakes that we see in the data, the combination of low wealth and a low risk-aversion coefficient implies the same behaviour as high wealth and a high risk-aversion coefficient. Unrestricted estimation of the CRRA model does not allow us to identify W and γ separately. The wealth level of $W = 50,000$ is representative for household wealth in the Netherlands.¹⁶ Hence, if the participants of LINGO are representative for the Dutch population, a relative risk aversion parameter of $\gamma = 6.99$ would be our best guess.

The results displayed in Table 6 are reminiscent of the findings in the literature. In general, the literature has found that different measures of wealth have major effects for asset pricing models. For example, Campbell (1966) finds that, if the measure of wealth is increased by including human wealth, the estimate for the CRRA coefficient increases substantially. Jagannathan and Wang (1996) study a cross-section of stock returns and find that the estimates of the conditional capital asset pricing model (CAPM) are substantially improved by including a measure of human capital in investors' wealth.

Parameters a_1 and a_2 are the estimates of the 'guilder certainty equivalent'

¹⁶ Bloemen (1997) estimates median household wealth in the Netherlands at approximately $f50,000$ in 1993 and 1994. He also provides median household wealth estimates for various age classes and education levels of the head of the household. For the set of finals for which the age and the education level of both players were available (68 finals in total), we attributed to each player the median household wealth level for the age/education group to which the player belongs. The average of the household wealth levels thus obtained is approximately $f32,500$. Since this is only a very rough wealth estimate of the population of LINGO players, in the sequel we use $f50,000$ as our baseline wealth level. We note that, if the average Dutch household consists of two persons, then the wealth level per individual would be $f25,000$. A team participating in LINGO would again hold $f50,000$. Finally, we note that we have estimated our model on the subsample of finals for which we have the team-specific wealth estimates. However, we failed to obtain sensible results. With differences in wealth, the values of the comeback options should be allowed to vary across the teams. The size of the subsample, already small, does not admit such an extension.

values of coming back when playing the first and second final, respectively. That is, a_1 (a_2) guilders is the amount of money that players would be willing to pay to be allowed to come back after the first (second) final. The estimates for a_1 and a_2 are both highly significant and sizable, indicating that the possibility to come back plays an important role in the decisions taken. Note also that $a_1 > a_2$, as required by the model.

The results for the CARA utility specification in Table 7 are very similar to those for CRRA utility. The risk-aversion parameter is significantly different from zero (risk neutrality). With $W = 50,000$, its estimate of 0.12 implies an estimated degree of relative risk aversion of 6.0, which is somewhat smaller than what we obtained for CRRA utility. The option values and the fit are almost identical to those for the CRRA model. Without data on individual wealth, the two specifications have similar implications for behaviour in the LINGO game.

To give an impression of what the estimates imply for an agent's risk attitude, Table 8 considers the choice between x with certainty or a lottery with payout $(2x, a; p, 1 - p)$. That is, a lottery with pay-off $2x$ with probability p and some fixed amount a with probability $1 - p$. The table shows that an agent with constant relative risk aversion, no reference wealth ($W = 0$ and thus $\gamma = 0.42$) would only be willing to take the gamble if the probability of winning is at least 0.67. For the power utility function ($W = 0$), the critical probability does not depend on the stake x . Models with either $W = 10,000$ or $W = 50,000$ fit the data better, and imply that teams are more willing to play if the stakes are low. The critical probabilities are almost identical for the CRRA and CARA models.

A positive option value a reduces the critical probabilities. Therefore, first-time finalists will always play as long as the stakes are below a . Even with stakes larger than a , first-time finalists should almost always continue. Only if the stakes are above $f4,000$, do they require a survival probability that is more than one half. This usually does not happen until the fourth round of a final. The

Table 7
Constant Absolute Risk Aversion

	γ	a_1	a_2	σ	N	$\ln L$
All games	0.12 (0.01)	1,887 (208)	1,523 (236)	0.21	1,186	-351.48
Third finals	0.11 (0.03)			0.17	111	-50.48
NOT third finals	0.13 (0.02)	2,050 (230)	1,670 (262)	0.23	1,075	-299.31

Notes: Dependent variable is the play/stop decision (1/0). γ is the coefficient of absolute risk aversion in the CARA model. a_1 and a_2 are the equivalent guilder amounts of the 'comeback' utilities. $\ln L$ is the value of the loglikelihood at the optimum, and σ is the standard deviation of the error term. Standard errors are in parentheses. 'All games' uses all data, while 'Third finals' is only based on teams that are in the final for the third and last time. 'NOT third finals' is the complementary subsample. The null hypothesis of equality of γ in both subsamples cannot be rejected.

Table 8
Critical Win Probabilities (p^)*

γ	Parameters		Stakes (x)				
	Wealth	a	500	1,000	2,000	4,000	8,000
<i>CRR</i>							
0.42	0	0	0.67	0.67	0.67	0.67	0.67
2.06	10,000	0	0.52	0.55	0.59	0.65	0.73
6.99	50,000	0	0.52	0.53	0.57	0.63	0.73
0.42	0	1,801	0	0	0.11	0.43	0.54
2.06	10,000	1,896	0	0	0.06	0.45	0.63
6.99	50,000	1,890	0	0	0.06	0.44	0.64
<i>CARA</i>							
0.12	–	0	0.52	0.53	0.56	0.62	0.72
0.12	–	1,887	0	0	0.06	0.43	0.64

Notes: The entries report the critical probabilities p^* for various parameter combinations of the CRR and CARA utility functions and for various stakes x . The parameter a is the option value for teams that have the option to return in a next show. Parameter estimates are from Tables 6 and 7. The critical probabilities are defined as $p^* = [U(x) - U(a)]/[U(2x) - U(a)]$ and denote the lowest win probability at which a player will prefer the lottery with payout $(2x, a; p, 1 - p)$ over the sure amount x .

most informative data are therefore the decisions made by third time finalists, who cannot return.

For this reason, we split the complete sample into a subsample containing third finals only and a subsample containing first and second finals only. The lower panels of Table 6 and 7 present the subsample estimates for CRR and CARA utility, respectively. For the third finals subsample the option values a_1 and a_2 are, of course, not identified. The point estimates of γ for the subsamples are not significantly different from each other. The estimate of γ is significant in both subsamples.

6. Model Extensions and Robustness Checks

The results presented in Section 5 for the basic expected-utility model suggest substantial risk aversion on the side of the LINGO players.¹⁷ In this section, we address the robustness of these results by estimating extensions of the basic expected-utility model and providing some additional checks on these models.

6.1. Heterogeneity in Ability

We found that the amount a team wins in the prefinal, W_0 , had a significant effect on the frequency distribution of the number of balls to be drawn in a final round (Table 2). This is most likely the result of differences in the ability to guess words. Heterogeneity in ability means that some teams have a higher

¹⁷ The model estimated in the previous section will from now on be referred as the ‘basic expected-utility model’ to distinguish it from the extensions discussed in this section.

survival probability in the final and, therefore, play longer. Better teams will also have a higher option value of returning. They can expect to win more next show, and might even have a higher probability of winning the prefinal in the next show. To test these effects, we split the sample of finals into those for which $W_0 \leq 400$ and those for which $W_0 > 400$, and allow a_1 and a_2 to differ between these two subgroups. For each of the subgroups, we compute the survival probabilities according to (1), replacing the overall frequency distribution f_j with the corresponding subgroup-specific frequency distribution.

Table 9 presents the results for CARA utility. The new estimate for γ , 0.13, hardly differs from the original estimate. The estimates of a_1 and a_2 for the ‘high-ability’ group exceed those for the ‘low-ability’ group, although the difference is only significant in the case of a_2 . The results for CRRA utility are not reported, because they almost exactly coincide with those for CARA utility.

6.2. Flow Utility

The participants in the show may not be a random sample from the Dutch population, but they may rather be people who enjoy these types of games or who like to be on television. To investigate this issue empirically, suppose that players derive some additional utility purely from the joy of participating in a television show. That is, we assume that they have the augmented utility function

$$\tilde{U}(x, i) = U(x) + c_1 i \tag{15}$$

where i is the round of play in a LINGO final. For teams that stop, the last round of play is $(i - 1)$. When $c_1 > 0$, the additional term makes playing another round more attractive. The critical probability p^* now becomes

Table 9
Model Extensions

	γ	$a_{1,low}$	$a_{1,high}$	$a_{2,low}$	$a_{2,high}$	c_1	μ	ω	σ	$\ln L$
Het. ability	0.13 (0.01)	1,805 (224)	2,041 (246)	1,173 (265)	2,047 (337)				0.22	-348.12
Flow utility	0.20 (0.07)	1,492 (280)		1,127 (293)		0.41 (0.17)			0.25	-347.50
Decision Weights	0.24 (0.07)	1,590 (300)		1,200 (320)			-0.51 (0.26)	0.76 (0.30)	0.23	-347.42

Notes: Dependent variable is the play/stop decision (1/0). γ is the coefficient of absolute risk aversion in the CARA model. a_1 and a_2 are the equivalent (guilder) amounts of the ‘comeback’ utilities; ‘low’ refers to the subgroup that enters the final with $W_0 \leq 400$ while ‘high’ refers to the subgroup that enters the final with $W_0 > 400$. $\ln L$ is the value of the loglikelihood at the optimum, and σ is the standard deviation of the error term. Standard errors are in parentheses. The row ‘Het. ability’ (heterogeneity in ability) refers to case in which a_1 and a_2 are allowed to differ across the subgroups with $W_0 \leq 400$ and $W_0 > 400$. The row ‘Flow utility’ refers to the model where players derive utility from every round of play. For the ‘Decision Weights’ model probabilities are transformed to decision weights according to the double LOGIT transformation (17) with parameters μ and ω . See Tables 6 and 7 for further explanatory notes.

$$p^* = \frac{U(x) - U(a) - c_1}{U(2x) - U(a)} \quad (16)$$

For CRRA and CARA utility, it is straightforward to see that the critical probabilities are increasing over the rounds in a given final if $c_1 > 0$ (use Lemma 1 and observe that $c_1/[U(2x) - U(a)]$ is decreasing in x , $x > a$). Hence, as before, the multi-stage decision problem of the players is reduced to a single, elementary lottery. The row labelled 'Flow utility' in Table 9 contains the estimation results. We present them only for CARA utility, because the results for CRRA utility are very similar. Parameter c_1 is significantly positive. Compared to the basic model, it is now more attractive to continue longer, *ceteris paribus*. Hence, to explain that players occasionally do stop, the estimate for γ increases to 0.20. The difference in the loglikelihoods of the regressions (see also Table 7) indicates that the basic model is rejected.

6.3. Decision Weights

A third possible extension of the basic model is motivated by empirical evidence based on psychological models of behaviour which suggests that individuals tend to be overconfident (Tversky and Kahneman, 1974). In the context of LINGO, such overconfidence would lead players to overestimate their abilities in solving word puzzles. To model systematic deviations between actual probabilities and the weights that subjects give to alternative outcomes, (cumulative) prospect theory introduces the concept of decision weights (Kahneman and Tversky, 1979, 1982; Tversky and Kahneman, 1992; Benartzi and Thaler, 1995). If people are overconfident, the decision weight attached to survival is larger than the actual survival probability. In that case people will tend to play longer than would be justified by their risk preferences. However, the longer people play in the LINGO final, the lower our estimates of risk aversion. When we mistakenly model the decision rule using the actual survival probabilities, we will therefore obtain biased low estimates of the risk-aversion parameter.

We will employ the following flexible double Logit transformation from actual probabilities p to decision weights $\pi(p)$:

$$\pi(p) = \frac{e^q}{1 + e^q}$$

$$q = \frac{\ln[p/(1-p)] - \mu}{\omega} \quad (17)$$

If $\mu = 0$ and $\omega = 1$, the transformation reduces to $\pi = p$. In this specification, μ acts primarily as a shift parameter and allows for a systematic over- or underestimation of the actual survival probability. Parameter ω determines the curvature of the function $\pi(p)$.

Players now base their decisions on their own subjective assessment of their survival chances. Fortunately, since $\pi(p)$ is a monotonic and continuous function of the actual probabilities, the players' decision problem simplifies

again to a single elementary lottery, where the actual survival probability p_i is now replaced by the decision weight $\pi(p_i)$. As under the cumulative prospect Theory, we assume that the decision weights sum to one. Hence, the decision weight attached to not surviving this round is given by $1 - \pi(p_i)$.

In the probit model, the actual probabilities p_i are replaced by the decision weights $\pi(p_i)$,

$$y_i^* = \pi(p_i) - p_i^* + \epsilon_i \quad (18)$$

while p_i^* is still the critical probability defined in (3). Separate identification of the risk aversion parameter and the parameters in the transformation from actual probabilities to decision weights rests on the functional specifications imposed on utility and on the transformation. The different nonlinearities of $\pi(p_i)$ and p_i^* identify the parameters. Different assumptions about utility could, in principle, affect the estimated risk aversion and the estimated probability transformation. However, in the following, we stick to CRRA or CARA utility, which we have employed throughout.

In Table 9, we report the parameter estimates for CARA utility. The estimate of the risk aversion γ has doubled compared to the basic expected utility model. At the same time, the estimated decision weights are higher than the actual probabilities. The bias is illustrated in Fig. 5, which depicts the estimated function $\pi(p)$. For every observation i , we have that $\pi(p_i) > p_i$ and, in many cases, the decision weight is substantially larger than the actual probability. If

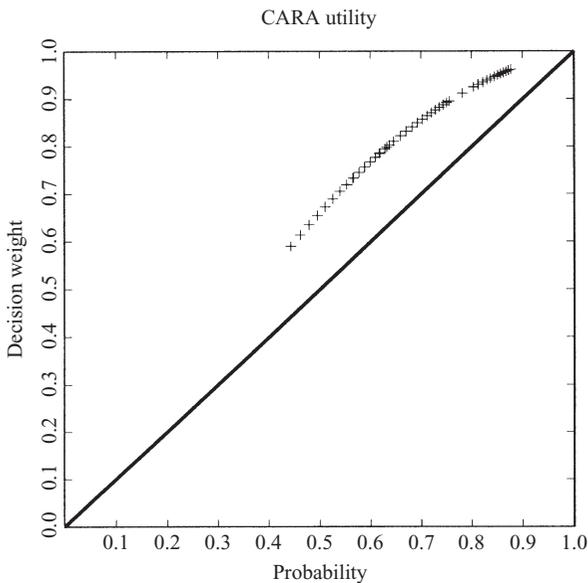


Fig. 5. *Empirical Decision Weights*

This figure shows the estimated relation between actual win probabilities and decision weights for the estimated parameters of the CARA utility function combined with (17).

The plus signs (+) correspond to the observed survival probabilities.

that is the case, a higher value for γ is needed to explain the observed stop decisions in the data. The value of the log-likelihood functions indicates that the model fits the data significantly better than the basic model.¹⁸

6.4. Variable Addition Tests

In Section 3 a rule-of-thumb model was estimated to describe the behaviour of teams. If utility maximisation provides a correct description of players' decisions, the expected-utility model should be able to encompass the rule-of-thumb specification. For this purpose, we estimate the probit model

$$y_i^* = \alpha(p_i - p_i^*) + \beta z_i + \epsilon_i \quad (19)$$

where $p_i - p_i^*$ is the difference between the actual and the critical survival probability computed at the maximum likelihood estimate for the CARA specification in Table 7 (CRRA utility gives virtually the same results). If the theoretical choice model provides an adequate description of the behaviour of the players, all explanatory variables in the vector z_i in (19) should have zero coefficients.

The first row of Table 10 shows that the basic expected-utility model certainly beats the most naïve decision rule, ie that people always play: the constant term is not significant. The second row indicates that the expected-utility model correctly captures the effects of the survival probability (*BAD*,

Table 10
Diagnostic Tests

Model	Instruments	Lik.-Ratio
Expected utility	$C, p_i - p_i^*$	0.48
Expected utility	$C, p_i - p_i^*, STAKE, PROB, BAD, BALLS, FINAL$	7.71
Expected utility	All	38.0*
Heterogeneous ability	All	28.0*
Utility flow	All	44.1*
Decision weights	All	37.3*

Notes: The third column reports the likelihood ratio tests for adding the 'rules-of-thumb' variables to the different theoretical models mentioned in the first column. 'Instruments' denotes the set of variables included on the right-hand side of (19). 'All' are $(p_i - p_i^*)$, a constant term (C) and the set of explanatory variables in Section 4: *STAKE, FINAL, BALLS, BAD, LAST, W₀* and *PROB*. The critical probability p_i^* is computed using the CARA parameter estimates in Tables 7 and 9. Dependent variable is the play/stop decision (1/0). The different models are the basic expected-utility model in Section 5, the heterogeneous ability extension in Section 6.1, the utility flow extension in Section 6.2, and the decision weights model in Section 6.3. Degrees of freedom for the likelihood ratio test with 'All' instruments is 9. An asterisk (*) denotes significance at the 1% level.

¹⁸ Empirical results for CRRA (not reported in the table) are qualitatively similar. With wealth equal to $f50,000$, the estimate of relative risk aversion almost doubles to 13.5, while the estimated decision weights are identical to those of the CARA model.

BALLS, *PROB*), the ‘comeback’ options (through *FINAL*) and the size of the lotteries (*STAKE*). When we use the full set of instruments, we find that *LAST* denoting the number of the round of the final in which the decision is made – and W_0 are both significant.

In each of the models, *LAST* enters with a positive coefficient, which indicates that people are more inclined to play the further the final progresses, *ceteris paribus*. We include *LAST* to test for heterogeneity in risk preferences. More risk-averse players will stop earlier, and, therefore, in later rounds of a given (first, second or third) final, the population of players will, on average, be less risk averse. At given odds, such players are more likely to continue.

Another way to pick up the effect of potential heterogeneity in risk aversion is to split the complete sample into a subsample containing third finals only and a subsample containing first and second finals only. People who enter the final for the third day in a row must have lost on the two previous days. This could select only those people who are relatively less risk averse than the average player. However, although the estimates of γ for third finalists in Tables 6 and 7 are indeed lower than those for the other subsample, the difference is not significant.

Similarly, teams with a higher initial stake W_0 tend to continue longer than implied by the expected-utility model and the extensions we considered. W_0 is even significant for the model with heterogeneity in ability. An explanation might be that players that enter with a high W_0 are relatively more overconfident. Hence, despite that, on average, they continue longer because of better survival chances, this may be strengthened further because they overestimate their odds to a larger extent. However, the data do not support this explanation. When we estimate the decision weights model for the two subsamples based on W_0 , the estimates of the parameters in (17) are not significantly different between the subsamples.

6.5. Simulation

By simulating the model, we can compare additional implications of the estimated decision rule with the actual data from the game. This provides an extra check on the estimated decision rule. We simulate the model by repeatedly drawing an initial stake W_0 from its empirical distribution. Starting with W_0 , the first round of the game is played. We simulate how many and which balls players draw from the urn, which determines whether they lose or proceed to the second round. After updating the survival probability, the estimated decision rule determines whether players continue or stop. The procedure continues until players stop, lose, or win all five rounds of the final.¹⁹

Simulation is most straightforward for third finals, since players have no possibility of returning in future shows. Repeating the simulation 500,000

¹⁹ A more detailed description is available on request.

times, we record the distribution of the gains in the final (W_1) and the percentage of *STOPPERS*, *LOSERS* and *WINNERS*. Simulation results for the third final using the basic expected-utility model are given in Table 11. The simulated data closely corresponds with the actual distribution of *STOPPERS*, *LOSERS* and *WINNERS* in the sample. The simulated gains are higher than the actual W_1 , although well within a two-standard-error range. It appears that our model does not systematically make wrong predictions about the circumstances in which players will stop or play.

The most interesting result from the simulations is that they allow us to compute the option value a_2 endogenously. When players lose in a second final, they are allowed to come back in one more show. From the average utility of the simulated money gains, we find the certainty equivalent $a_2 = 1,402$. Analogously, for the sample of losers in the second final, the certainty equivalent of the average utility of their money gains gives an estimate of a_2 equal to 1,423, which is very close to the simulated value. Both are less than the estimated parameter value $\hat{a}_2 = 1,523$ in Table 7, which means that players either overestimate their winning chances or attach value to being allowed to appear in a television game one more time.²⁰

7. Conclusions

One of the main problems in economics is the assessment of individuals' risk attitudes. In this paper, we have used a new and unique dataset, obtained from the Dutch television game show LINGO, to estimate the degree of risk aversion of its players. LINGO provides an excellent natural experiment for this purpose: the players' decisions are simple (stop or continue), the dataset is large and the money stakes are high.

Our main findings are the following. First, using our basic expected-utility framework, we found clear evidence of risk aversion (although less than what is needed to explain the equity-premium puzzle), both for our complete

Table 11
Simulation of Third Finals

	<i>WINNERS</i>	<i>STOPPERS</i>	<i>LOSERS</i>	Prize money (W_1)	Implied option (a_2)
Simulation	7%	34%	59%	2,982	1,402
Actual data	5%	39%	55%	2,760	1,423

Notes: Simulation results are averages based on 500,000 simulations of a third final, assuming CARA utility with $\gamma = 0.12168$. The actual percentage of *WINNERS*, *STOPPERS* and *LOSERS* is taken from the last column of Table 3. The actual money gains (W_1) are for third finalists. The calculation of the implied option value is available upon request.

²⁰ All the results above are for the CARA utility functions. Results for CRRA utility are very similar and not reported.

sample of finalists and for third finalists only (ie, the group that might be expected to be the least risk averse). Based on our initial wealth estimate of $f50,000$, we estimate the coefficient of relative risk aversion at approximately 7. The implied willingness to engage in an elementary lottery depends on the size of the stake: the required minimum probability of winning in a lottery in which the stake is either doubled or everything is lost rises from 53% for a stake of $f1,000$ to 73% for a stake of $f8,000$.

The second main finding is that, for plausible generalisations of our basic framework, the estimated degree of risk aversion increases. One extension is to allow for the players' utility not only to depend on the money gains but also on the act of playing itself (e.g., players just enjoy playing these types of games or they like to be on television).²¹ Another extension is to allow decisions to be based on decision weights rather than the actual probabilities. In accordance with empirical findings in other contexts, the estimates suggest that players display a strong tendency to overestimate their chances of success. This could most likely be attributed to an overestimation of the ability to solve word puzzles. To explain the stop decisions in the data with this model, the estimated risk aversion needs to increase.

University of Amsterdam and CEPR

LIFE/Maastricht University and CEPR

Date of receipt of first submission: May 1999

Date of receipt of final typescript: February 2001

Appendix: Additional Details of the Game

Clues for Guessing the Word: Whenever a couple has to guess a new word, two letters (and their positions in the word) are already given. If some attempt yields one or more additional letters that feature in the word, then these letters are also indicated. It is also indicated whether the position of these letters within the word is correct or not. Using this information, words can usually be found in a small number of attempts.

The jackpot: In the pre-final, there is also a 'jackpot', which increases by $f50$ after each correctly guessed word. After each word they solve, a couple can draw two balls, without replacement, from an urn containing green and red balls (this is a different urn from the one used in the final). The jackpot accumulates over the shows until some couple has drawn a total of three green balls in a given pre-final. This couple receives the amount in the jackpot, which starts at zero again. Thus, the starting value of the jackpot in some given pre-final is the jackpot value at the end of the previous pre-final. Gains from the jackpot are not at risk in the final. However, when the jackpot has risen to a high value, this may provide an incentive to continue longer in the final, because the failure to survive this final (if it is not the third final) yields another attempt at the jackpot in the next pre-final.

²¹ Another reason, not investigated here, why our baseline risk-aversion estimates may understate players' true risk aversion is the 'gambling with the house money effect' (Thaler and Johnson, 1990): players are not yet accustomed to the money they have won so far and are, therefore, more willing to bet their stakes.

References

- Arrow, K. J. (1982). 'Risk perception in psychology and economics', *Economic Inquiry*, vol. 20, pp. 1–9.
- Barksey, R. B., Juster, F. T., Kimball, M. S. and Shapiro, M. D. (1997). 'Preference parameters and behavioral heterogeneity: an experimental approach in the health and retirement study', *Quarterly Journal of Economics*, vol. 112, pp. 537–79.
- Benartzi, S. and Thaler, R. H. (1995). 'Myopic loss aversion and the equity premium puzzle', *Quarterly Journal of Economics*, vol. 110, pp. 73–92.
- Bloemen, H. G. (1997). 'A description of the distribution of wealth in the Netherlands, using microdata', Working paper, Tilburg University.
- Camerer, C. F. (1987). 'Do biases in probability judgment matter in markets? Experimental evidence', *American Economic Review*, vol. 77, pp. 981–97.
- Campbell, J. Y. (1996). 'Understanding risk and return', *Journal of Political Economy*, vol. 104, pp. 298–345.
- Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.
- Conlisk, J. (1996). 'Why bounded rationality?', *Journal of Economic Literature*, vol. 34, pp. 660–700.
- Friend, I. and Blume, M. E. (1975). 'The demand for risky assets', *American Economic Review*, vol. 65, pp. 900–22.
- Gertner, R. (1993). 'Game shows and economic behavior: Risk taking on "card sharks"', *Quarterly Journal of Economics*, vol. 108, pp. 507–21.
- Gneezy, U. (1997). 'Essays in behavioral economics', Dissertation, Tilburg University.
- Hartog, J., Ferrer-i-Carbonell, A. and Jonker, N. (2000). 'On a simple measure of risk aversion', Working paper, University of Amsterdam.
- Jagannathan, R. and Wang, Z. (1996). 'The conditional CAPM and the cross section of expected returns', *Journal of Finance*, vol. 51, pp. 3–53.
- Kachelmeier, S. J. and Shehata, M. (1992). 'Examining risk preferences under high monetary incentives: experimental evidence from the People's Republic of China', *American Economic Review*, vol. 82, pp. 1120–41.
- Kahneman, D. and Tversky, A. (1979). 'Prospect theory: An analysis of decision under risk', *Econometrica*, vol. 47, pp. 263–91.
- Kahneman, D. and Tversky, A. (1982). 'The psychology of preferences', *Scientific American*, vol. 246, pp. 167–73.
- Kocherlakota, N. R. (1996). 'The equity premium: It's still a puzzle', *Journal of Economic Literature*, vol. 34, pp. 42–71.
- Mankiw, N. G. and Zeldes, S. P. (1991). 'The consumption of stockholders and non-stockholders', *Journal of Financial Economics*, vol. 29, pp. 97–112.
- Mehra, R. and Prescott, E. C. (1985). 'The equity premium puzzle', *Journal of Monetary Economics*, vol. 15, pp. 145–61.
- Metrick, A. (1995). 'A natural experiment in "jeopardy"', *American Economic Review*, vol. 85, pp. 240–53.
- Rabin, M. (2000). 'Diminishing marginal utility of wealth cannot explain risk aversion', in (D. Kahneman and Tversky, A. eds) *Choices, Values, and Frames*, Cambridge: Cambridge University Press.
- Thaler, R. H. and Johnson, E. J. (1990). 'Gambling with the house money and trying to break even: The effects of prior outcomes on risky choice', *Management Science*, vol. 36, pp. 643–60.
- Tversky, A. and Kahneman, D. (1974). 'Judgement under uncertainty: Heuristics and biases', *Science*, vol. 185, pp. 1124–31.
- Tversky, A. and Kahneman, D. (1992). 'Advances in prospect theory: Cumulative representation of uncertainty', *Journal of Risk and Uncertainty*, vol. 5, pp. 297–323.
- Wakker, P. P. (1994). 'Separating marginal utility and probabilistic risk aversion', *Journal of Risk and Uncertainty*, vol. 36, pp. 1–44.