# High-Stakes Failures of Backward Induction

Bouke Klein Teeselink[1], Dennie van Dolder[2],
Martijn J. van den Assem[3,4] and Jason D. Dana[5]

[1] *Department of Political Economy, King's College London*
[2] *Department of Economics, University of Essex*
[3] *School of Business and Economics, Vrije Universiteit (VU) Amsterdam*
[4] *Tinbergen Institute*
[5] *Yale School of Management*

January 3, 2024

**Abstract**

We examine high-stakes strategic choice using more than 40 years of data from the American TV game show *The Price Is Right*. In every episode, contestants play the *Showcase Showdown*, a sequential game of perfect information for which the optimal strategy can be found through backward induction. We find that contestants systematically deviate from the subgame perfect Nash equilibrium. These departures from optimality are well explained by a modified agent quantal response model that allows for limited foresight. The results suggest that many contestants simplify the decision problem by adopting a myopic representation and optimize their chances of beating the next contestant only. In line with learning, contestants' choices improve over the course of our sample period.

**Keywords**: backward induction; limited foresight; omission bias; quantal response equilibrium; subgame perfect Nash equilibrium
**JEL Classifications**: C72; D01; D91

# 1 Introduction

Many economic interactions are sequential in nature. A negotiator who makes a bargaining offer, an entrepreneur who considers whether to enter a market, and a corporate manager who decides how many goods to produce, all need to consider the subsequent actions of others. Such situations can be modeled as finite sequential games of perfect information, for which the subgame perfect Nash equilibria can be found through backward induction (von Stackelberg, 1934; Selten, 1978; Dixit, 1982; Rubinstein, 1982).

Unfortunately, the descriptive accuracy of game-theoretic equilibria is difficult to test in the field, because agents' choice options, payoffs, and the information they have are normally not (or not straightforwardly) observable. When choices deviate from equilibrium play, it consequently remains unclear whether the behavior is truly suboptimal or whether the deviations are the result of incorrect modeling assumptions. To avoid this joint-hypothesis problem, tests of equilibrium play typically rely on laboratory experiments in which all factors are perfectly controlled. Experimental work generally finds that people often deviate from the equilibrium strategies, casting doubt on the descriptive validity of backward induction as a solution concept (Rosenthal, 1981; McKelvey and Palfrey, 1992; Fey et al., 1996; Binmore et al., 2002; Johnson et al., 2002; Levitt et al., 2011; Dufwenberg and Van Essen, 2018). The generalizability of experimental findings to real-world situations, however, is subject to debate (Binmore, 1999; Levitt and List, 2007*a*,*b*; Falk and Heckman, 2009; Camerer, 2015). Critics argue that it is not very surprising that experimental subjects frequently fail to adopt equilibrium strategies, most notably because of low incentives and little experience with the task.

The present paper examines the optimality of strategic decisions in the quasi-controlled high-stakes field setting of the *Showcase Showdown* (SCSD), a finite sequential game of perfect information that is played twice in every episode of the long-running American TV game show *The Price Is Right*. In this game, described in more detail in Section 2, three contestants take turns to spin a wheel that contains all multiples of 5 in the range 5–100.[1] Immediately after spinning the wheel once, the contestant has to decide whether to spin the wheel again. Their score is the outcome of the first spin if they spin only once, and the sum of the two spin outcomes if they spin twice. The contestant whose score is closest to 100 without going over wins the game and proceeds to the so-called *Showcase* final, where they compete with the winner of another SCSD to win a set of prizes worth tens of thousands of

[1]Henceforth we refer to the contestant who spins first as Contestant 1, to the contestant who spins second as Contestant 2, and to the contestant who spins last as Contestant 3.

dollars in expectation. If their score is exactly 100 they win one or two cash bonus prizes on top of their qualification for the final.

To make the optimal choice, the contestants thus need to weigh the possibility of obtaining a more competitive score and having a shot at the bonus prizes against the risk of self-elimination. Coe and Butterworth (1995), Grosjean (1998), and Tenorio and Cason (2002) derive the unique subgame perfect Nash equilibrium (SPNE) for this game for various combinations of bonus prizes and expected showcase values. The three contestants' equilibrium strategies, which can be found through backward induction, take the form of decision rules that dictate whether a contestant should stop or use their second spin.

The characteristics of the SCSD make it an appealing test bed for assessing the descriptive validity of backward induction as a solution concept in the field. First, as in carefully designed lab experiments, the task is well-defined and both the choice options and the choice-relevant information that is available to contestants are known. Second, the prizes that can be won dwarf the payoffs that are typically employed in experiments. Third, the SCSD has been repeated numerous times under similar conditions, creating the opportunity of a large-scale statistical analysis. Other benefits of this long history are that contestants can be expected to be familiar with the game and that we can explore potential learning effects.[2]

At the same time, the game show setting may evoke external validity concerns because of selection procedures and because of the unusual conditions under which choices are made. Section 5 reflects on these concerns. Any possible downside, however, should be evaluated in the light of the availability of better alternatives. Other opportunities for a large-scale, high-stakes field test of backward induction are incredibly scarce, if not absent. Hence, following List (2023), the unique setting should be embraced and not dismissed for its idiosyncrasies.

The present paper is not the first to use a TV game show as a real-world naturally-occurring laboratory. Game shows have been used to study a wide range of other topics in economics, such as decision making under risk (Gertner, 1993; Metrick, 1995; Post et al., 2008; Bombardini and Trebbi, 2012), discrimination (Levitt, 2004; Belot et al., 2010), bargaining (van Dolder et al., 2015), willingness to compete (Hogarth et al., 2012; Buser et al., 2023), giving (Eberhardt et al., 2024), and cooperation (List, 2006; Oberholzer-Gee et al., 2010; van den Assem et al., 2012; Turmunkh et al., 2019).

We analyze a large sample of 10,071 renditions of the SCSD. In every rendition,

[2]The SCSD has also been proposed as a useful classroom tool for teaching probability and game theory (Burks and Jaye, 2012; Swenson, 2015).

three contestants make a spin decision, but a substantial fraction of the 30,213 decisions are trivial and of little value to our study. For Contestant 2 and 3, decisions are trivial when their first-spin outcome is lower than the best preceding score (in which case they always spin again). For Contestant 3, who spins last, decisions are also trivial when their first-spin outcome is higher than the best preceding score (in which case they always stop).[3] We omit such decisions from our empirical analysis, and exclusively focus on the decisions of Contestant 1 and on the remaining decisions of Contestant 2.

We start our analysis by examining whether, when, and how contestants deviate from the SPNE. We find that Contestants 1 and 2 frequently make suboptimal decisions, and that the error rate of Contestant 1 is somewhat higher than that of Contestant 2. Strikingly, Contestant 1 almost exclusively errs by underspinning: stopping when it is optimal to spin. Contestant 2's mistakes, by contrast, are considerably more symmetric and involve only slightly more underspinning than overspinning.

We then consider several explanations for suboptimal play that are well-rooted in the literature. First, we examine whether contestants depart from the equilibrium strategy because they make random errors in evaluating the expected utility of their two choice options, and expect others to make similar mistakes. To test this explanation, we estimate an agent quantal response equilibrium model (AQRE; McKelvey and Palfrey, 1998). We find that a substantial proportion of the deviations from the SPNE can be explained by random evaluation errors. The decisions of Contestant 2 are largely consistent with the model's probabilistic predictions, but the model fails to capture most of the underspinning of Contestant 1.

Next, we consider the possible role of omission bias, which is the tendency to favor harmful inactions over harmful actions (Ritov and Baron, 1990, 1992; Spranca et al., 1991; Feldman et al., 2020). Systematic underspinning in the SCSD can be explained by a preference for elimination after not spinning (by an opponent who obtains a higher score) over elimination after spinning (by exceeding 100 points).[4] We find that allowing for omission bias in the AQRE model improves the goodness-of-fit for Contestant 1, but at the same time introduces systematic prediction errors for Contestant 2. Hence, omission bias fails to adequately explain the behavior that we observe.

[3] Contestant 3 faces a nontrivial decision when they tie with the best preceding score, but such situations are relatively rare.

[4] Walker et al. (2018) propose the related concept of sudden death aversion: the tendency to avoid strategies that can lead to immediate defeat, even if these are optimal. In our setting, sudden death aversion and omission bias are indistinguishable, because spinning (acting) entails the risk of immediate elimination whereas not spinning (not acting) postpones possible elimination.

Another possible explanation is that contestants may not properly backward induct, and instead adopt a simplified representation of the game. Prior research suggests that people have limited foresight, and look only one or a few steps ahead in multi-stage strategic situations (Jehiel, 1995, 1998, 2001; Johnson et al., 2002; Gabaix and Laibson, 2005; Gabaix et al., 2006; Mantovani, 2016; Ke, 2019; Rampal, 2022; Baranski and Reuben, 2023). We adjust our baseline AQRE model to allow for the possibility that a contestant myopically behaves as if the next stage of the game is also the last. Such a simplified frame lowers Contestant 1's propensity to spin, because beating only one subsequent contestant in expectation requires a lower score than beating two. For Contestant 2, limited foresight coincides with full backward induction because the next stage is also the last stage of the game. Our limited foresight model accurately describes the observed behavior of contestants. According to the estimation results, approximately 41 percent of the contestants simplify the game by looking only one step ahead.

The overall conclusion therefore is that the deviations from the SPNE in this high-stakes game are well explained by a combination of random evaluation errors and limited foresight, and that the role of omission bias is limited. This conclusion is robust to various alternative modeling assumptions.

Our findings diverge markedly from those of Tenorio and Cason (2002), who explore a relatively small sample of renditions of the SCSD from 1994 and 1995. They conclude that omission bias is a plausible explanation for the deviations from the SPNE in their data, but their analysis is limited to simple comparisons, and their evidence derives almost exclusively from decisions by Contestant 1 (due to a lack of informative observations for Contestant 2). The present paper uses a considerably larger sample, with many informative observations for both Contestant 1 and Contestant 2, which allows for the estimation of structural decision models and for tests of competing hypotheses.

Our results are striking in the light of the long history of the show and its popularity. A natural question is whether contestants' behavior converges towards the SPNE over time. To answer this question, we subdivide our sample into four periods. For Contestant 1, we find that the frequency of deviations from the SPNE decreases substantially and monotonically. For Contestant 2, there is no clear trend over time. When we estimate the limited foresight model, we find that the fraction of spinning decisions that are made in accordance with limited foresight decreases monotonically from 65 percent in the first period to 18 percent in the last. Despite this strong improvement, the results show that many contestants remain unable to follow the optimal strategies deriving from backward induction, even after several decades of *The Price Is Right*.

The remainder of the paper is structured as follows. Section 2 introduces the game show and the SCSD in more detail and outlines the equilibrium strategies. Section 3 discusses the data and provides a descriptive analysis of deviations from equilibrium play. Section 4 presents the main analyses and results, various robustness checks, and the learning analysis. Section 5 concludes and discusses our findings.

# 2 The Game and Its Equilibrium Strategies

*The Price Is Right* was first aired in the United States in 1956. Through the years, the format was introduced in many other countries, but here we exclusively consider the American version. Every episode consists of multiple games. The game that is central in our paper—the *Showcase Showdown* (SCSD)—was first included in 1975. Apart from a change in 2008 (see below), the SCSD has remained the same since 1979. We exclusively consider episodes from 1979 onwards.

Every episode contains two renditions of the SCSD, with three contestants each. Prior to the SCSD, every contestant plays two other games: the so-called *One Bid* game, and a "pricing game". In the *One Bid* game, four contestants guess the retail price of a consumer product (such as a microwave or television).[5] The contestant whose guess is closest to the actual retail price without going over wins the product, gets to play one of the many different pricing games, and will be one of the SCSD contestants.[6] In their pricing game, the contestant can win one or more prizes, often by guessing the retail prices of consumer goods. After three contestants have won a *One Bid* game and completed their pricing game, the first SCSD is played. In the next part of every episode this combination of three *One Bid* games, three pricing games, and one SCSD is repeated.

The winners of the two SCSDs proceed to the final of the episode. In this so-called *Showcase* round, the two finalists have to guess the retail price of their own respective showcase, which typically consists of multiple valuable prizes such as a car, furniture, electronics, or a trip. The contestant whose guess is closest to the retail price without exceeding it wins the contents of their showcase. If the winner's guess is within a specified range below the retail price ($100 until 1997-98, $250

[5] Contestants are selected by the producers through interviews with ticketed audience members shortly before to the recording of the episode.

[6] Bennett and Hickman (1993), Berk et al. (1996), and Healy and Noussair (2004) use the *One Bid* game to study strategic decision making. Atanasov et al. (2023) use it to study own-gender favoritism.

from 1998-99 onwards) they win both showcases; if both finalists' guesses exceed the retail price both showcases remain unclaimed.

In the SCSD, our game of interest, three contestants take turns to spin a big wheel that contains all multiples of 5 up to 100. The contestant with the lowest (highest) prior winnings spins first (last). Immediately after observing the outcome of their first spin, a contestant has to decide whether to spin the wheel again.[7] Their score is the outcome of the first spin if they spin once, and the sum of the two spin outcomes if they spin twice. The contestant whose score is closest to 100 without exceeding is the winner and proceeds to the *Showcase* round.[8] If two or three contestants tie for the highest score, they enter a "spin-off" in which each of them spins the wheel once more; the one who spins the highest number is the winner. This procedure is repeated in the case of further ties.

On top of securing a place in the lucrative final, SCSD contestants can also win one or two monetary bonus prizes. If a contestant scores exactly 100 points, they receive \$1,000 plus a bonus spin that yields an additional \$10,000 (\$5,000 before 2008-09) if the wheel lands on 5 or 15, or \$25,000 (\$10,000 before 2008-09) if it stops at 100. If two or three contestants tie at a score of 100, the outcome of their bonus spin counts as their spin-off score.

The optimal strategy for a contestant depends on the expected showcase value and the bonus prizes. Coe and Butterworth (1995), Grosjean (1998), and Tenorio and Cason (2002) derive the SPNE for a limited set of combinations of these values. The three contestants' equilibrium strategies, which can be found through backward induction, take the form of optimal stopping rules that dictate when a contestant should not use their second spin. Our sample covers a large time span, over which the average retail price of the showcases varied considerably, and during which there was a change in the bonus prizes. We therefore derive the optimal stopping thresholds for a large set of combinations of expected showcase values and the two bonus schemes.

In line with previous work, we assume (i) that spin outcomes follow a discrete uniform distribution from 5 to 100 with steps of 5, (ii) that contestants are risk neutral, and (iii) that the chance of winning the *Showcase* round after winning the SCSD is 50 percent. Section 4.4 examines the sensitivity of our results to the last two assumptions. We use numerical methods to compute the optimal strategies.[9]

[7] The wheel must be spun for at least one full revolution.

[8] If the third contestant beats the best preceding score with their first spin, or if the first two contestants went over 100, the third contestant automatically advances to the *Showcase* round. In the latter case, Contestant 3 does spin the wheel once to try to win a bonus prize by spinning exactly 100, but they are not given the choice to spin a second time.

[9] As also noted by Tenorio and Cason (2002), a complete analytical solution is infeasible due to the discrete partitions of the wheel, the possibility of ties, and the presence of bonuses.

Table 1 shows each contestant's optimal strategy for various ranges of expected showcase values, denoted $E(S)$, and for the two different bonus schemes. For brevity, the table displays the optimal strategies for empirically relevant ranges of $E(S)$ only.[10] Furthermore, it omits the trivial optimal decisions of Contestant 2 and Contestant 3 in situations where their first-spin outcome is lower than the best preceding score (where they should always spin again), and that of Contestant 3 when their first-spin outcome beats the best preceding score (where they are automatically declared the winner). The optimal stopping rule jumps discretely because the values on the wheel are multiples of five.

Consider, for example, a rendition of the SCSD with the most recent bonus scheme and where $E(S) = \$25,000$. Contestant 3 faces a nontrivial decision only when they tie the best preceding score. If they tie with one previous contestant, they should stop when the tie is at 55 or more (and spin otherwise). In the case of a three-way tie, the stopping threshold is 70.

The optimal strategies of the other two contestants can be found through backward induction. Assuming that Contestant 3 strictly adopts the optimal approach, Contestant 2 is best off by stopping at 60 or more if they beat the score of Contestant 1, and by spinning otherwise. In the case of a tie with Contestant 1, Contestant 2's stopping threshold is 70. Contestant 1 has to anticipate the decisions of Contestant 2 and 3. Assuming that these two both follow the optimal strategy, Contestant 1's stopping threshold is 70.

# 3 Data and Preliminary Results

Our data are from the *The Price Is Right Episode Guide*.[11] We accessed this fan-edited website on 21 June 2021. At that time, it contained $5,834$ detailed recaps of episodes of *The Price Is Right* from 1979 onwards. We successfully scraped the data for one or both SCSDs for 5,307 episodes. After omitting special episodes with a deviating prize structure, and the one available episode from the 1978-79 season, our final sample contains 10,071 SCSDs from 5,235 different episodes that were aired between 1979-80 and 2020-21.[12] In most cases, we additionally obtained contestants' names, the accumulated value of the prizes they earned prior to the SCSD, and the

[10] When $E(S)$ goes to zero, the optimal stopping threshold converges to 100.

[11] See https://tpirepguide.com.

[12] Some types of special episodes featured a deviating SCSD bonus scheme or extra-valuable prizes in the *Showcase* round. We identified and omitted such episodes using https://www.priceisright.fandom.com, a collaborative website dedicated to *The Price Is Right*. We omit the one episode from the 1978-79 season because we cannot reliably estimate the expected showcase value for that season.

Table 1: Optimal strategies

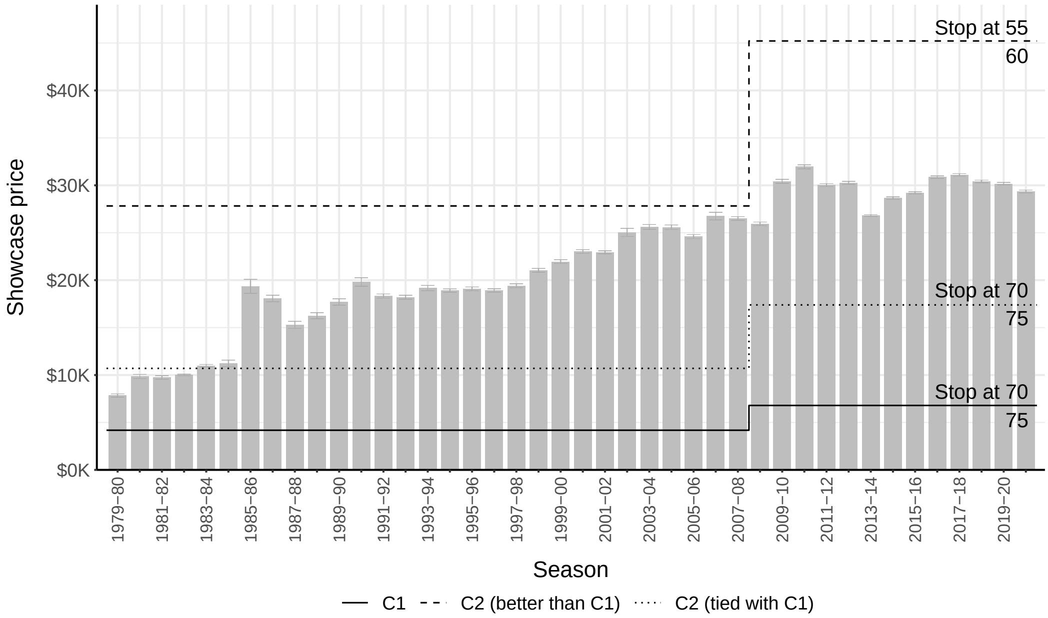
| Contestant | First spin | $E(S)$ | Stopping threshold |
|---|---|---|---|
| **Panel A: Bonus Scheme 1 (until 2008-09)** | | | |
| C1 | | $\$1,532 \leq E(S) < \$4,180$ | 75 |
| C1 | | $E(S) \geq \$4,180$ | 70 |
| C2 | Better than C1 | $\$2,564 \leq E(S) < \$27,826$ | 60 |
| C2 | Better than C1 | $E(S) \geq \$27,826$ | 55 |
| C2 | Tied with C1 | $\$2,503 \leq E(S) < \$10,702$ | 75 |
| C2 | Tied with C1 | $E(S) \geq \$10,702$ | 70 |
| C3 | Tied with C1 or C2 | $\$2,000 \leq E(S) < \$4,000$ | 60 |
| C3 | Tied with C1 or C2 | $E(S) \geq \$4,000$ | 55 |
| C3 | Tied with C1 and C2 | $\$2,400 \leq E(S) < \$6,000$ | 75 |
| C3 | Tied with C1 and C2 | $E(S) \geq \$6,000$ | 70 |
| **Panel B: Bonus Scheme 2 (from 2008-09 onwards)** | | | |
| C1 | | $\$2,489 \leq E(S) < \$6,792$ | 75 |
| C1 | | $E(S) \geq \$6,792$ | 70 |
| C2 | Better than C1 | $\$4,167 \leq E(S) < \$45,217$ | 60 |
| C2 | Better than C1 | $E(S) \geq \$45,217$ | 55 |
| C2 | Tied with C1 | $\$4,068 \leq E(S) < \$17,391$ | 75 |
| C2 | Tied with C1 | $E(S) \geq \$17,391$ | 70 |
| C3 | Tied with C1 or C2 | $\$3,250 \leq E(S) < \$6,500$ | 60 |
| C3 | Tied with C1 or C2 | $E(S) \geq \$6,500$ | 55 |
| C3 | Tied with C1 and C2 | $\$3,900 \leq E(S) < \$9,750$ | 75 |
| C3 | Tied with C1 and C2 | $E(S) \geq \$9,750$ | 70 |

*Notes:* The table shows the optimal strategies for various ranges of expected showcase values and for the two different bonus schemes. Under Bonus Scheme 1 (Panel A), the bonus prizes are \$1,000, \$5,000, and \$10,000; under Bonus Scheme 2 (Panel B), the bonus prizes are \$1,000, \$10,000, and \$25,000. The first column indicates whether the contestant is the first (C1), second (C2), or third (C3) to spin. The second column indicates whether the contestant's first spin beats or ties the best preceding score. The third column gives the range for the expected showcase value. The last column gives the optimal stopping threshold: the first-spin outcome at or above which the contestant should stop, and below which they should spin again. The table omits the trivial optimal decision of Contestant 2 and Contestant 3 in situations where their first-spin outcome is lower than the best preceding score (always spin), and that of Contestant 3 when their first-spin outcome beats the best preceding score (always stop).

stated retail prices of the showcases. Table A1 in the Appendix shows the numbers of episodes, SCSDs, and stated showcase prices in our sample for every season.

As a first analysis, we explore the extent to which contestants' spinning decisions are consistent with the SPNE. Because almost all Contestant 3's decisions are trivial—they are automatically declared the winner if their first-spin outcome beats the best preceding score, and by default spin again if it is lower—we focus exclusively on Contestants 1 and 2. For the same reason, we omit decisions of Contestant 2 that follow first-spin outcomes that are lower than the score of Contestant 1. This leaves us with 10,071 spinning decisions for Contestant 1 and 4,488 for Contestant

Figure 1: Average stated retail price of showcases across seasons




*Notes:* The figure shows the average stated retail price of showcases for every season. Error bars depict standard errors around the mean. Horizontal lines indicate the most relevant expected showcase values at which the optimal stopping thresholds change. Table A1 in the Appendix shows the number of included showcases per season.

2.

The previous section showed how the optimal stopping rule depends on a contestant's assessment of the expected showcase value. We make the simplifying assumption that this subjective value equals the average stated retail price of the showcases in the given season, and examine the sensitivity of our results to this assumption in Section 4.4.

Figure 1 shows the average stated retail price per season. Throughout our sample period, this average increased from $7,838 (1979-80) to $29,342 (2020-21), or by approximately 3.3 percent per year. For comparison, the inflation in the US over this period was 3.0 percent per year (US Consumer Price Index; OECD, 2021). The horizontal lines indicate the most relevant thresholds at which the optimal stopping rules change. The jumps reflect the change of the bonus prizes. At any expectation higher than these thresholds, the stopping rules remain the same.

For Contestant 1, the average retail price was always well above the threshold values of $\$4,180$ (until 2008-09) and $\$6,792$ (from 2008-09 onwards). Hence, throughout our entire sample period Contestant 1 optimizes their play by stopping if and only if their first spin is 70 or higher. For Contestant 2 we need to distinguish between situations where their first spin beats the score of Contestant 1, and situations where they tie.[13] Contestant 2's optimal stopping rule in situations where their first spin beats Contestant 1's score was also constant over time: the average retail price never exceeded the critical values of $27,826 (until 2008-09) and $45,217 (from 2008-09 onwards), which means that they should stop if and only if their first spin is 60 or higher. For ties the optimal stopping rule did change. Most of the time—from the 1983-84 season onward—Contestant 2 was best off by stopping if and only if the tie was at 70 or higher. Until the start of the 1983-84 season the stopping threshold was 75.

When we compare contestants' actual decisions with the optimal decisions, we observe that only a small proportion deviate. For Contestant 1, 93.4 percent of the 10,071 decisions are in accordance with the equilibrium strategy. For Contestant 2, 95.9 percent of the 4,488 decisions are optimal. The low rates at which contestants depart from optimality are not very surprising, because most decisions are easy. When we exclusively consider "difficult" choice situations—which we define as situations where the first-spin outcome is no more than two steps below the stopping threshold and no more than one step above it—we find that 72.9 percent of the 2,069 decisions of Contestant 1 and 79.5 percent of the 790 decisions of Contestant 2 are in accordance with the equilibrium strategy. Hence, for these more difficult choice situations, the rates of departure from optimality are considerable.

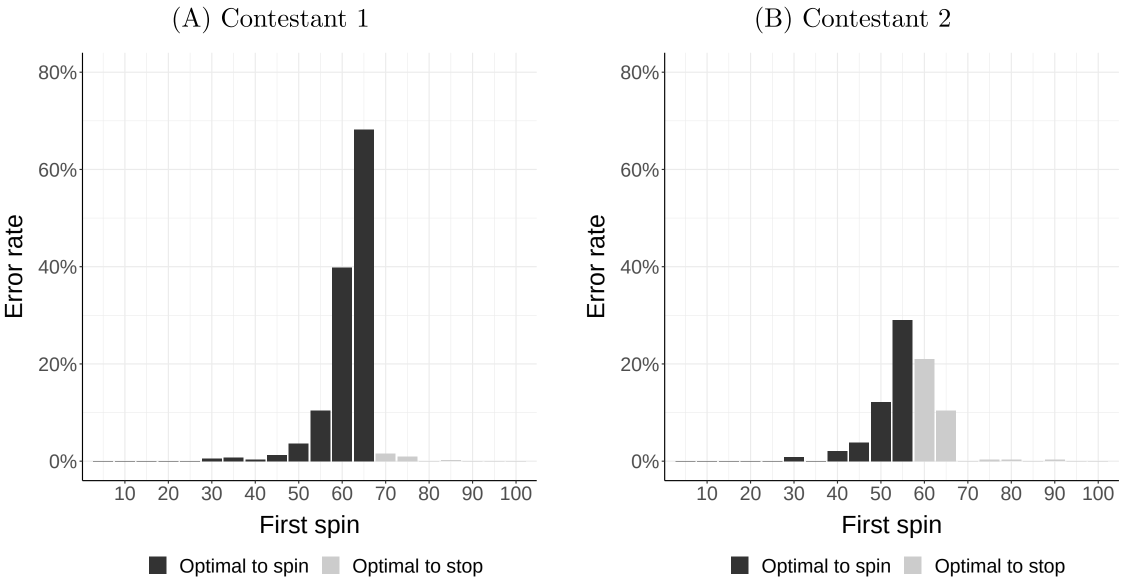
The deviations tend to be in one direction. If Contestants 1 and 2 were to follow the optimal strategy, they would spin in 65.9 and 32.4 percent of all situations in our sample, respectively. In reality, however, they spin only 59.6 and 30.9 percent of the times. For the more difficult situations, the optimal spinning rates are 49.1 and 46.3 percent, whereas the actual rates are only 23.3 and 39.7 percent. Hence, these global spinning statistics indicate that there is systematic underspinning, especially for Contestant 1.

Figure 2 shows how often Contestant 1 (Panel A) and Contestant 2 (Panel B) deviate from the optimal strategy, for every possible first-spin outcome.[14] The dark grey bars represent the deviations in situations where it is optimal to spin, the light grey bars show the deviations in situations where it is optimal to stop. Clearly, at first-spin outcomes of 60 and 65 Contestant 1 frequently departs from the equilibrium strategy. In these situations, spinning is optimal but many instead choose to stop. In contrast to these many underspinning errors, Contestant 1 displays hardly any overspinning errors. For Contestant 2, the pattern looks different. Contestant 2

[13]Ties are relatively rare. Out of the 4,488 nontrivial spinning choices that we have for Contestant 2, only 384 (8.6%) are ties.

[14]Figure 2 omits the (relatively rare) choice situations of Contestant 2 where they are tied with Contestant 1, because the optimal stopping threshold is different for these situations.

Figure 2: Deviations from the SPNE




*Notes:* The figure shows how often the decisions of Contestant 1 (Panel A, N=10,071) and Contestant 2 (Panel B, N=4,104) deviate from the optimal strategy, for every possible first-spin outcome. Panel B omits ties and thus exclusively considers choice situations where the first-spin outcome of Contestant 2 beats the score of Contestant 1. Dark gray bars depict first-spin outcomes at which it is optimal to spin, light gray bars depict first-spin outcomes at which it is optimal to stop.

departs less frequently from the optimal strategy than Contestant 1. Moreover, in comparison with Contestant 1, their deviations from optimality are considerably more symmetric.

The costs of deviating from the SPNE are often sizable. Table A2 in the Appendix shows the costs when the expected showcase value is $30,000 and the second bonus scheme applies. This combination is representative for the last 12 years of our sample period.[15] For example, for Contestant 1, stopping after a first-spin outcome of 65 lowers their expected earnings by $393, and stopping after 60 lowers it by $1,189. For Contestant 2, spinning after beating Contestant 1 with a first spin of 60 lowers their expected earnings by $1,008. The costs of stopping at 55—Contestant 2's most frequent deviation from the SPNE—are relatively small at only $55.

[15] It is infeasible to provide a complete picture because the costs depend on the expected value of the showcase and the bonus prizes, which both change over the course of our sample period.

# 4 Analyses and Results

In the current section we propose and test three possible explanations for contestants' deviations from the SPNE. Section 4.1 introduces our baseline structural model, which allows for the possibility that contestants make random evaluation errors. Section 4.2 then extends this model with the possibility of omission bias, whereas Section 4.3 instead extends it to allow for the possibility of limited foresight. Section 4.4 presents several robustness checks, which include tests of various alternative explanations. Last, Section 4.5 exploits the longitudinal dimension of the data to explore whether there is evidence of learning over the years.

## 4.1 Random Errors

The SPNE is based on the assumption that contestants perfectly maximize their expected utility, and never make mistakes. In reality, people of course will make mistakes. In the SCSD, the costs of mistakes vary between choice situations, and strongly depend on a contestant's first spin outcome. In situations where spinning is only slightly better than stopping, or vice versa, even a small evaluation error could lead a contestant to deviate from the optimal choice. Depending on the relative costs of over- or underspinning across all choice situations, random evaluation errors can lead to a pattern of systematic deviation from the SPNE.

Moreover, a player who realizes that the choices of their opponents are not flawless should take this into account in determining their optimal strategy. Factoring in the mistakes of others may lead to optimal strategies that differ from the SPNE (Goeree and Holt, 2001; Goeree et al., 2002, 2003). In the SCSD, mistakes of subsequent opponents generally lower the incentive to spin again. Therefore, in theory, the anticipation of mistakes could explain the underspinning as compared to the SPNE.

To examine the role of random errors, we adopt the Quantal Response Equilibrium (QRE) concept (McKelvey and Palfrey, 1995; Chen et al., 1997). The QRE is a stochastic generalization of the Nash equilibrium, and commonly used to account for bounded rationality in strategic settings (see, for example, Capra et al., 1999; Anderson et al., 2001; Goeree et al., 2002, 2003; Moinas and Pouget, 2013; Goeree et al., 2016, 2017). The main underlying idea is that people make random mistakes in evaluating the expected utilities of choice alternatives, and that they anticipate that others do the same. Because the SCSD is a sequential game, we consider the Agent Quantal Response Equilibrium (AQRE), a modification of the QRE for extensive-form games (McKelvey and Palfrey, 1998). The AQRE concept has found many applications (see, for example, Fey et al., 1996; McKelvey and Palfrey, 1998; Deck, 2001; Cason and Reynolds, 2005; Cai and Wang, 2006; McKelvey and Patty, 2006; Fehr et al., 2021).

Almost all of Contestant 3's decisions are trivial, and therefore we assume that Contestants 1 and 2 expect Contestant 3 to play their SPNE strategy without error. Similarly, we assume that Contestant 1 does not expect Contestant 2 to err after a first-spin outcome that is worse than Contestant 1's score, because Contestant 2 by default always spins again in such situations.

For all nontrivial choice situations, let $EU_{ij}^{s}(\cdot)$ denote the expected utility of action $s \in \{Spin, Stop\}$ for Contestant $i \in \{1, 2\}$ in SCSD $j \in \{1, 2, \ldots, J\}$. Contestants make random evaluation errors $\varepsilon_{ij}^{s}$ and mistakenly consider $\widehat{EU}_{ij}^{s}(\cdot) = EU_{ij}^{s}(\cdot) + \varepsilon_{ij}^{s}$. Following convention, we assume that $\varepsilon_{ij}^{s}$ is independently and identically distributed according to an extreme value distribution, which leads to the following predicted spin probabilities (Goeree et al., 2005; Haile et al., 2008; Goeree et al., 2020):

$$P_{ij}^{Spin} = \frac{e^{\lambda EU_{ij}^{\text{Spin}}}}{e^{\lambda EU_{ij}^{\text{Spin}}} + e^{\lambda EU_{ij}^{\text{Stop}}}} \tag{1}$$

$\lambda$ can be interpreted as contestants' rationality or payoff sensitivity parameter. If $\lambda \to 0$, contestants make completely random choices and spin with a 50 percent likelihood; if $\lambda \to \infty$, they follow the payoff-maximizing strategy with certainty. In Section 4.4, we consider a more flexible specification that allows $\lambda$ to differ between Contestant 1 and Contestant 2.

The expected utilities of spinning and stopping both depend on the resulting probability of winning the SCSD, the shape of the utility function, the chance of winning the showcase after winning the SCSD, and the showcase value; for spinning, the expected utility in addition depends on the bonus prizes. In our main analyses we assume risk neutrality. We also assume that contestants believe that they have a 50 percent chance of winning the showcase after winning the SCSD, and that the expected showcase value equals the average stated retail price of all showcases in the entire running season. We examine the sensitivity of the results to these assumptions in Section 4.4.

We convert all nominal monetary values to 2015 dollars using the US Consumer Price Index (OECD, 2021).[16] To obtain more readable coefficients, we divide all monetary values by 1,000. The next subsections expand this baseline model with additional parameters that capture omission bias and limited foresight. We use

[16]For completeness, Section 4.4 also gives the results without correcting for inflation.

Table 2: Estimation results

| | SPNE | | Baseline | | Omission bias | | Limited foresight | | OB & LF | |
|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | - | - | 1.384 | (0.028) | 1.508 | (0.032) | 1.579 | (0.039) | 1.583 | (0.039) |
| $\gamma$ | - | - | - | - | 0.803 | (0.050) | - | - | 0.110 | (0.088) |
| $\beta$ | - | - | - | - | - | - | 0.410 | (0.024) | 0.371 | (0.039) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -1,964 | | -1,843 | | -1,794 | | -1,794 | |
| AIC | - | | 3,930 | | 3,689 | | 3,593 | | 3,593 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.959 | 0.934 | 0.959 | 0.952 | 0.943 | 0.953 | 0.959 | 0.953 | 0.945 |
| Hit rate (difficult) | 0.729 | 0.795 | 0.729 | 0.795 | 0.815 | 0.703 | 0.817 | 0.795 | 0.817 | 0.713 |
| Brier score | 0.066 | 0.041 | 0.046 | 0.035 | 0.037 | 0.037 | 0.034 | 0.035 | 0.034 | 0.035 |
| Brier score (difficult) | 0.271 | 0.205 | 0.176 | 0.173 | 0.132 | 0.185 | 0.120 | 0.174 | 0.120 | 0.174 |
| Spinning bias | -0.063 | -0.014 | -0.040 | 0.001 | -0.013 | 0.021 | 0.007 | 0.001 | 0.006 | 0.004 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.217 | 0.009 | -0.104 | 0.112 | 0.003 | 0.014 | -0.001 | 0.028 |

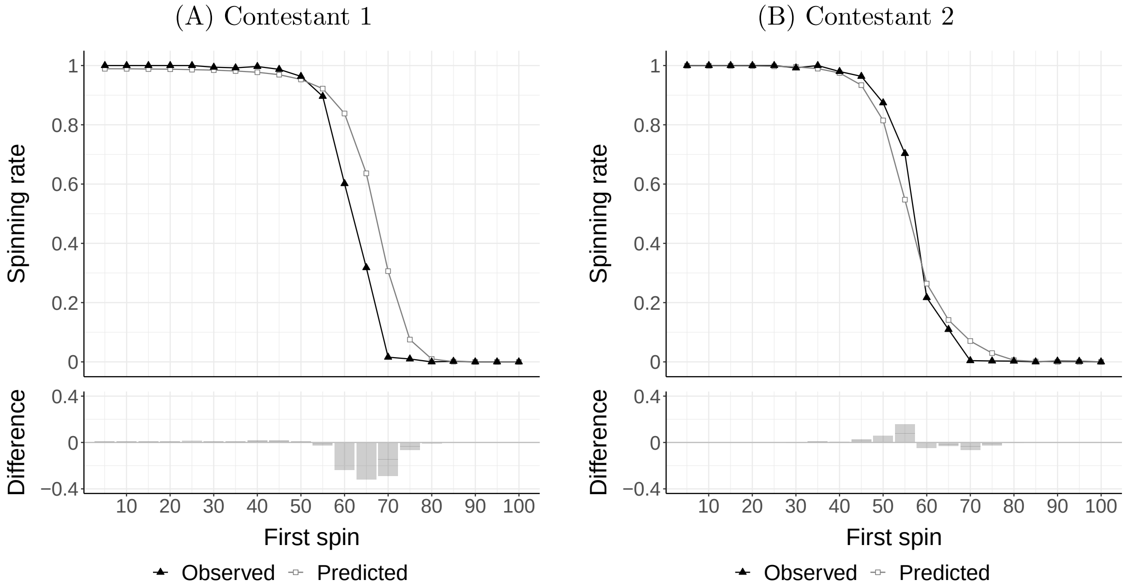
*Notes:* The table shows the estimated parameters and goodness-of-fit statistics of various structural models of strategic decision making. *SPNE* is the model that adopts the binary predictions from the subgame perfect Nash equilibrium, *Baseline* is the baseline AQRE model, *Omission bias* is the AQRE model that incorporates omission bias, *Limited foresight* is the AQRE model that allows for limited foresight, and *OB & LF* is the AQRE model with both omission bias and limited foresight. $\lambda$ is the estimated rationality parameter, $\gamma$ is the estimated disutility of self-elimination, and $\beta$ is the estimated probability of limited foresight. Standard errors are in parentheses. $N$ is the number of spinning decisions, *Log likelihood* is the log likelihood value of the estimation, and *AIC* is the Akaike Information Criterion value. Other goodness-of-fit measures are given separately for Contestant 1 (*C1*) and Contestant 2 (*C2*), both for all choice situations combined and for relatively difficult choice situations only. Difficult choice situations are choices where the first-spin outcome is no more than two steps below the stopping threshold and no more than one step above it. *Hit rate* is the fraction of correctly predicted decisions, *Brier score* is the mean squared prediction error, and *Spinning bias* is the average difference between contestants' actual spinning decisions and the model's spinning predictions.

maximum likelihood techniques to estimate the parameters. Because the number of decisions of Contestant 1 in our sample is more than twice the number of decisions of Contestant 2, we weigh the observations of Contestant 1 by $(N_1+N_2)/2N_1$ and those of Contestant 2 by $(N_1 + N_2)/2N_2$, such that the overall weights for the two types of contestants are equal and the average weight across all individual contestants remains unity. Without this weighting, the results would be disproportionately driven by the choices and idiosyncrasies of Contestant 1.

Table 2 presents the results. To compare how well the baseline AQRE model explains contestants' behavior relative to the SPNE, we consider three goodness-of-fit statistics: the hit rate, the Brier score, and the spinning bias.

The hit rate of the model is the fraction of correctly predicted decisions. A prediction is defined as correct if the model assigns a 50 percent or greater probability to the contestant's actual decision. The baseline model correctly predicts 93.4 percent of Contestant 1's decisions and 95.9 percent of Contestant 2's decisions. These high hit rates are not surprising, because most decisions in our sample are easy. For relatively difficult choice situations—where the first-spin outcome is no more than two steps below the stopping threshold and no more than one step above it —the hit rate of the baseline model is 72.9 percent for Contestant 1 and 79.5 percent for

Figure 3: Empirical spinning rates and baseline model predictions




*Notes:* The figure shows the observed spinning rate and the average probabilistic prediction of the baseline model for Contestant 1 (Panel A, N=10,071) and for Contestant 2 (Panel B, N=4,488) for all possible first-spin outcomes. The lower parts of the panels show the differences between the observed and predicted spinning rates.

Contestant 2. These hit rates are identical to those for the SPNE, suggesting that allowing for evaluation errors does not add any descriptive power. Due to the binary nature of "hits", however, the measure is rather crude. In contrast to the SPNE, the predictions of the AQRE are probabilistic, and much of the variation in these probabilities is not reflected in the hit rate.

To assess the difference between the observed choices and the probabilistic predictions, we calculate the Brier score (Brier, 1950). The Brier score is the mean squared prediction error. For the binary predictions of the SPNE, the Brier score is the complement of the hit rate. Compared to the Brier scores for the SPNE, those for the baseline model are substantially lower. The improvement is especially strong for Contestant 1; for difficult decisions, for example, the statistic declines from 0.271 to 0.176.

The Brier score is a good measure to assess overall predictive accuracy, but it is uninformative of the degree to which the model systematically over- or underpredicts contestants' propensity to spin. To visually explore whether there is any systematic deviation, Figure 3 plots the actual spinning rates against the average probabilistic predictions of the baseline AQRE model for every possible first-spin outcome. The figure clearly shows that Contestant 1 underspins relative to the predictions. At first-spin outcomes of 60, 65 and 70, the fraction of contestants who actually use their second spin is approximately 25-30 percentage points lower than predicted. For Contestant 2 the differences are much smaller, with the actual spinning rate on average being slightly higher than predicted.

The spinning bias quantifies the degree of systematic deviation, and is calculated as the average difference between contestants' actual spinning decisions, which take a value of either 0 (stop) or 1 (spin), and the model's probabilistic spinning predictions, which can take any value between 0 and 1. A positive value of this goodness-of-fit statistic indicates overspinning, a negative value underspinning. Confirming the pattern in Figure 3, the spinning bias is negative for Contestant 1: -4.0 percentage points at the aggregate level, and -21.7 percentage points for the relatively difficult first-spin outcomes. This degree of contestants' systematic underspinning according to the baseline model is high, but lower than the negative spinning bias of Contestant 1 relative to the SPNE (-6.3 and -25.8 percentage points, respectively). For Contestant 2, the spinning bias is positive and close to zero: 0.1 percentage points across all choices, and 0.9 for the more difficult ones.

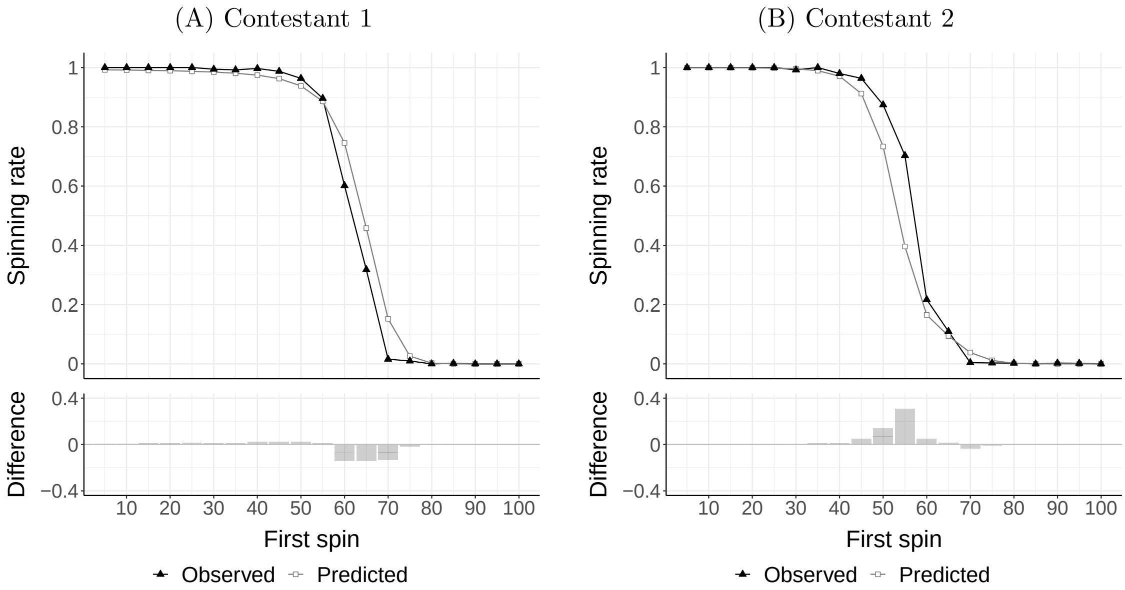
Taken together, these findings suggest that random evaluation errors can explain some of the deviations from the SPNE. The systematic underspinning of Contestant 1, however, remains largely unexplained.

## 4.2 Omission Bias

Tenorio and Cason (2002) explore a relatively small sample of renditions of the SCSD from 1994 and 1995, and similarly report that contestants tend to stop when it is actually optimal to spin. Their evidence derives primarily from Contestant 1, as their sample of informative decisions of Contestant 2 is too small to draw reliable conclusions. Tenorio and Cason propose that the underspinning can be explained by omission bias—the tendency to favor harmful inactions over harmful actions (Ritov and Baron, 1990, 1992; Spranca et al., 1991; Feldman et al., 2020). Other research shows that omission bias can play an important role in settings where decision makers face a choice between action and inaction. Examples include vaccination decisions, debt repayment, blackjack, and sports refereeing (Ritov and Baron, 1990; Asch et al., 1994; Carlin and Robinson, 2009; DiBonaventura and Chapman, 2008; Moskowitz and Wertheim, 2011; Hallsworth et al., 2023).

In the SCSD, contestants will be less likely to spin if they prefer elimination after not spinning (by an opponent who obtains a higher score) over elimination after spinning (by exceeding 100 points). To examine whether omission bias can explain the observed behavior, we extend the baseline structural model with $\gamma$, a parameter that captures the disutility of self-elimination. In our baseline specification, the

Figure 4: Empirical spinning rates and omission bias model predictions




*Notes:* The figure shows the observed spinning rate and the average probabilistic prediction of the omission bias model for Contestant 1 (Panel A, N=10,071) and for Contestant 2 (Panel B, N=4,488) for all possible first-spin outcomes. The lower parts of the panels show the differences between the observed and predicted spinning rates.

utility from elimination is always zero; in this alternative specification, however, the utility is $-\gamma$ if the contestant spins again and exceeds the maximum score of 100. The utility of elimination after an opponent obtains a higher score remains zero.

Table 2 shows the results for the AQRE model with omission bias. The estimated value of $\gamma$ is 0.803, implying that the disutility of losing through self-elimination is equivalent to the disutility of a monetary loss of \$803 (in 2015 dollars). This model explains contestants' choices better than the baseline model, also when we account for its additional parameter: both the log-likelihood and the AIC show a substantial improvement. A likelihood-ratio test confirms that the model with omission bias significantly outperforms the baseline model ($\chi^2(1) = 242.43$, $p < 0.001$).

The separate goodness-of-fit measures for the two contestants show that the omission bias model provides a better account of Contestant 1's decisions but a worse account of Contestant 2's decisions, as compared to the baseline model. For Contestant 1, the overall hit rate improves from 93.4 to 95.2 percent, and the hit rate for difficult decisions improves from 72.9 to 81.5 percent. The improved fit for Contestant 1 is also reflected in lower Brier scores. The opposite holds for Contestant 2: the overall hit rate deteriorates from 95.9 to 94.3 percent, the hit rate for more difficult decisions deteriorates from 79.5 to 70.3 percent, and the Brier scores increase.

Figure 4 compares the actual spinning rates and the average probabilistic predictions of the omission-bias model for all first-spin outcomes. For Contestant 1, as compared to the baseline model, the predictions are substantially closer to the actual spinning rates. As shown in Table 2, the remaining spinning bias of Contestant 1 is -1.3 percentage points, which compares favorably to the -4.0 percentage points of the baseline model. For relatively difficult first-spin outcomes the degree of underspinning decreases from 21.7 to 10.4 percentage points.

The reduction of the systematic prediction error for Contestant 1, however, is largely offset by an increase for Contestant 2. Contestant 2 clearly overspins relative to the predictions of the omission bias model. Their spinning bias increases from 0.1 to 2.1 percentage points across all choices, and from 0.9 to 11.2 percentage points for the more difficult ones.

Altogether, omission bias thus fails to adequately explain contestants' behavior. The additional parameter partially captures the underspinning of Contestant 1 and improves the overall fit of the model, but at the same time introduces large systematic prediction errors for Contestant 2.

## 4.3 Limited Foresight

A possible alternative explanation for the suboptimal behavior of contestants is limited foresight. A large body of theoretical research proposes that people reason only one or a few steps ahead (Jehiel, 1995, 1998, 2001; Jackson and Wolinsky, 1996; Gabaix and Laibson, 2005; Ke, 2019; Bossaerts et al., 2022; Rampal, 2022). Several experimental studies support this notion (Neelin et al., 1988; Johnson et al., 2002; Gabaix et al., 2006; Mantovani, 2016; Barreda-Tarrazona et al., 2021; Rampal, 2022; Baranski and Reuben, 2023).

To simplify the decision problem, an SCSD contestant may adopt a myopic representation and optimize their chances of beating the next contestant only. If Contestant 1 only considers Contestant 2 in their spinning choice and ignores the presence of Contestant 3, then Contestant 1 will be less inclined to spin because beating only one subsequent contestant in expectation requires a lower score than beating two. For Contestant 2, limited foresight coincides with full backward induction because the next stage is also the last stage of the game.

Our limited foresight model expands the baseline model with the possibility that contestants reduce complexity by considering the next contestant only. We assume that myopic contestants believe that this next contestant behaves as if they are the last. A contestant adopts the simplified frame with probability $\beta$, and correctly considers all future contestants with probability $1-\beta$. In this mixture model, the likelihood of spinning is the probability-weighted average of the likelihood under the assumption of limited foresight (with probability $\beta$) and the likelihood under full backward induction (with probability $1 - \beta$).

The penultimate column of Table 2 shows the results for the limited foresight model. The estimated $\beta$ coefficient is 0.410, suggesting that 41 percent of the spinning decisions are made in accordance with limited foresight, while the remaining 59 percent are consistent with full backward induction. The empirical fit is much better than the fit of the baseline and omission bias models: both the log-likelihood and the AIC show considerable improvements. A likelihood-ratio test confirms that the current model outperforms the baseline model ($\chi^2(1) = 338.99$, $p < 0.001$), and a Vuong test for non-nested models confirms that it also outperforms the omission bias model ($Z = 6.86$, $p < 0.001$).

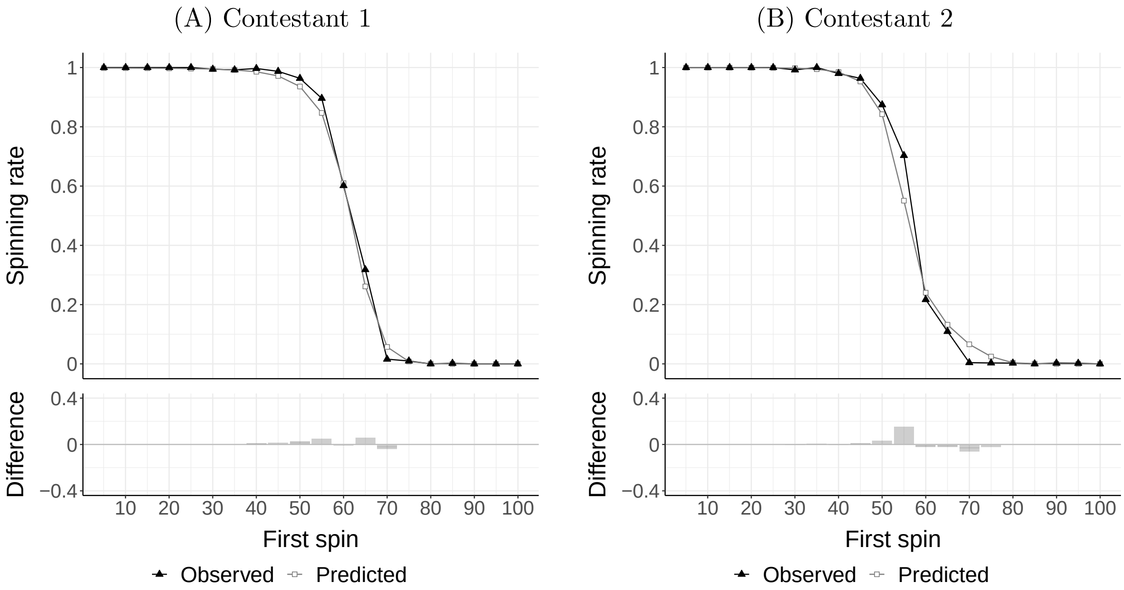
As compared to the omission bias model, the limited foresight model provides a slightly better account of Contestant 1's decisions, and a substantially better account of Contestant 2's decisions. For Contestant 1, the hit rates are nearly identical to those of the omission bias model, and the Brier scores are better. For Contestant 2, the overall hit rate increases from 94.3 to 95.9 percent, the hit rate for difficult first-spin outcomes increases from 70.3 to 79.5 percent, and the Brier scores improve.

Figure 5 plots the actual spinning rates against the average probabilistic predictions of the limited foresight model, and shows that the model accurately captures the observed behavior. For both Contestant 1 and Contestant 2, the actual and predicted spinning rates approximately coincide. As also shown in Table 2, barely any spinning bias remains. Across all choices, Contestant 1 spins 0.7 percentage points more often than predicted by the model, and Contestant 2 spins only 0.1 percentage points more often. For the more difficult choices, the spinning biases are a mere 0.3 and 1.4 percentage points, respectively.

The limited foresight model thus provides an accurate account of contestants' spinning decisions. To examine whether contestants' choices are in addition partly driven by omission bias, we estimate a model that allows for both omission bias and limited foresight. The final column of Table 2 shows the estimation results. The results clearly speak against omission bias as a possible driver. First, the omission bias parameter is relatively small and statistically insignificant. Second, the goodness-of-fit of the model is similar to that of the limited foresight model: neither the log-likelihood nor the AIC value improves, and a likelihood-ratio test does not reject the hypothesis that the two models explain spinning choices equally well ($\chi^2(1) = 1.56$, $p = 0.212$).

All in all, the conclusion from these analyses is that the behavior of contestants is

Figure 5: Empirical spinning rates and limited foresight model predictions




*Notes:* The figure shows the observed spinning rate and the average probabilistic prediction of the limited foresight model for Contestant 1 (Panel A, N=10,071) and for Contestant 2 (Panel B, N=4,488) for all possible first-spin outcomes. The lower parts of the panels show the differences between the observed and predicted spinning rates.

well described by an AQRE model with limited foresight, where all contestants make random evaluation errors and many simplify the decision problem by myopically considering the next stage of the game only.

## 4.4 Robustness Checks

The structural models require a variety of assumptions. In the present subsection, we explore the sensitivity of the results to risk aversion (Section 4.4.1), to beliefs about the expected value of winning the SCSD (Section 4.4.2), to the weight attached to the opportunity of winning bonus prizes (Section 4.4.3), and to various other, more minor aspects (Section 4.4.4).

### 4.4.1 Risk Aversion

The choice between spinning and stopping is essentially a choice between two risky prospects. In the analyses thus far we assumed that contestants are risk neutral. Here we explore the sensitivity of the results to the alternative assumption that contestants are risk averse. We now assume that they have a constant absolute risk aversion (CARA) utility function of the form $U(x) = 1 - \exp^{\theta x}$, where $x$ is the monetary value of the prospective winnings and $\theta$ is the risk-aversion coefficient.

Table 3: Estimation results under alternative modeling choices (1/2)

| | SPNE | | Baseline | | Omission bias | | Limited foresight | | OB & LF | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A: High risk aversion** | | | | | | | | | | |
| $\lambda$ | - | - | 9.503 | (0.191) | 10.811 | (0.231) | 11.234 | (0.286) | 11.395 | (0.286) |
| $\gamma$ | - | - | - | - | 0.147 | (0.007) | - | - | 0.049 | (0.012) |
| $\beta$ | - | - | - | - | - | - | 0.499 | (0.024) | 0.376 | (0.039) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -2,037 | | -1,833 | | -1,791 | | -1,783 | |
| AIC | - | | 4,076 | | 3,669 | | 3,587 | | 3,573 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.959 | 0.934 | 0.955 | 0.950 | 0.943 | 0.953 | 0.955 | 0.953 | 0.949 |
| Hit rate (difficult) | 0.729 | 0.795 | 0.729 | 0.789 | 0.806 | 0.722 | 0.817 | 0.789 | 0.817 | 0.753 |
| Brier score | 0.066 | 0.041 | 0.049 | 0.035 | 0.037 | 0.037 | 0.034 | 0.035 | 0.034 | 0.035 |
| Brier score (difficult) | 0.271 | 0.205 | 0.195 | 0.158 | 0.133 | 0.171 | 0.122 | 0.158 | 0.121 | 0.158 |
| Spinning bias | -0.063 | -0.014 | -0.048 | -0.005 | -0.014 | 0.021 | 0.009 | -0.005 | 0.007 | 0.004 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.252 | -0.016 | -0.103 | 0.118 | 0.013 | -0.011 | 0.000 | 0.037 |
| **Panel B: Discounted showcase value** | | | | | | | | | | |
| $\lambda$ | - | - | 2.609 | (0.052) | 2.972 | (0.064) | 3.070 | (0.078) | 3.119 | (0.077) |
| $\gamma$ | - | - | - | - | 0.537 | (0.026) | - | - | 0.202 | (0.045) |
| $\beta$ | - | - | - | - | - | - | 0.498 | (0.025) | 0.360 | (0.039) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -2,038 | | -1,832 | | -1,797 | | -1,787 | |
| AIC | - | | 4,078 | | 3,667 | | 3,598 | | 3,580 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.956 | 0.934 | 0.956 | 0.952 | 0.943 | 0.953 | 0.956 | 0.953 | 0.947 |
| Hit rate (difficult) | 0.729 | 0.789 | 0.729 | 0.789 | 0.815 | 0.718 | 0.817 | 0.789 | 0.817 | 0.739 |
| Brier score | 0.066 | 0.044 | 0.049 | 0.035 | 0.036 | 0.037 | 0.034 | 0.035 | 0.034 | 0.035 |
| Brier score (difficult) | 0.271 | 0.211 | 0.193 | 0.162 | 0.132 | 0.173 | 0.121 | 0.162 | 0.120 | 0.161 |
| Spinning bias | -0.063 | -0.018 | -0.047 | -0.006 | -0.013 | 0.020 | 0.009 | -0.006 | 0.006 | 0.004 |
| Spinning bias (difficult) | -0.258 | -0.076 | -0.251 | -0.020 | -0.101 | 0.115 | 0.012 | -0.015 | -0.003 | 0.038 |
| **Panel C: No bonus prizes** | | | | | | | | | | |
| $\lambda$ | - | - | 1.458 | (0.030) | 1.530 | (0.032) | 1.609 | (0.039) | 1.606 | (0.040) |
| $\gamma$ | - | - | - | - | 0.532 | (0.050) | - | - | -0.185 | (0.087) |
| $\beta$ | - | - | - | - | - | - | 0.320 | (0.024) | 0.384 | (0.039) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -1,912 | | -1,856 | | -1,804 | | -1,802 | |
| AIC | - | | 3,826 | | 3,716 | | 3,613 | | 3,610 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.943 | 0.934 | 0.943 | 0.953 | 0.943 | 0.953 | 0.943 | 0.953 | 0.947 |
| Hit rate (difficult) | 0.729 | 0.701 | 0.729 | 0.701 | 0.817 | 0.701 | 0.817 | 0.701 | 0.817 | 0.727 |
| Brier score | 0.066 | 0.057 | 0.042 | 0.035 | 0.037 | 0.037 | 0.034 | 0.036 | 0.034 | 0.035 |
| Brier score (difficult) | 0.271 | 0.299 | 0.160 | 0.171 | 0.133 | 0.183 | 0.119 | 0.172 | 0.120 | 0.170 |
| Spinning bias | -0.063 | 0.026 | -0.032 | 0.009 | -0.014 | 0.022 | 0.006 | 0.009 | 0.007 | 0.004 |
| Spinning bias (difficult) | -0.258 | 0.141 | -0.183 | 0.038 | -0.107 | 0.109 | -0.007 | 0.038 | 0.001 | 0.012 |

*Notes:* The table shows the results for three alternative modeling choices. Panel A shows the results under the assumption that contestants have CARA utility, with a certainty equivalent of $2,500 for a 50-50 lottery of winning $25,000 or $0. Panel B shows the results under the assumption that contestants value showcases at 50 percent of the stated retail price. Panel C shows the results under the assumption that contestants ignore the bonus prizes. Other definitions are as in Table 2.

We set $\theta$ such that the certainty equivalent of a 50-50 lottery of winning $25,000 or $0 is $2,500, reflecting a high degree of risk aversion. To obtain more readable coefficients, we scale the utility function such that the utility of $1,000 equals unity.

Table 3, Panel A presents the results. Introducing high risk aversion leaves the goodness-of-fit statistics of the binary predictions of the SPNE completely unchanged, strongly worsens the log-likelihood of the baseline model, and somewhat improves the fit of the models with omission bias, limited foresight, and the combination of these. Contestant 1 still underspins compared to the predictions of the baseline model, whereas Contestant 2 still overspins compared to the predictions of the omission bias model. The limited foresight model again provides an accurate account of contestants' spinning decisions, and displays a significantly better fit than the omission bias model (Vuong test: $Z = 5.02$, $p < 0.001$). We find similar results for more moderate degrees of risk aversion, for example when we set $\theta$ such that the certainty equivalent of the 50-50 lottery of winning \$25,000 or \$0 is \$5,000 or \$10,000 (see Table A3 in the Appendix). The main conclusions from the previous analyses thus do not hinge on the assumption of risk neutrality.[17]

#### 4.4.2 Discounting the Showcase Value

The optimal strategies and stochastic model predictions depend on the expected value of the showcase, $E(S)$, relative to the value of the monetary bonus prizes. For the main analyses, we assumed that the value of a showcase equals its stated retail price. The stated retail price is a natural and salient value, but in reality contestants will likely discount it. The showcase prizes are selected by the game show producers, not by the contestants themselves, and will mostly not align well with contestants' preferences.[18] As a robustness check, we re-estimate the structural models under the alternative assumption that contestants value showcases at 50 percent of the stated retail price.[19]

Table 3, Panel B presents the results. Discounting the showcase value leads to a worse fit of the baseline model and stronger evidence of underspinning. This is not surprising, because a lower expected showcase value increases the relative attractiveness of the bonus prizes, and thus increases the incentive to spin a second time. At the same time, limited foresight again provides a better account of contestants' spinning decisions than omission bias (Vuong test: $Z = 4.83$, $p < 0.001$).[20]

[17]Under strong risk aversion (Table 3, Panel A), adding the possibility of omission bias to the limited foresight model yields a statistically significant improvement of the empirical fit (LR test: $\chi^2(1) = 15.99$, $p < 0.001$). Under medium and low risk aversion (Table A3 in the Appendix), this improvement is much weaker or absent (LR test medium: $\chi^2(1) = 3.84$, $p = 0.050$; low: $\chi^2(1) = 0.93$, $p = 0.335$).

[18]Contestants should further discount the showcase value because of taxes. Although taxes are levied over both (monetary) bonus prizes and (generally non-monetary) showcase prizes, taxes generally make the showcase prizes comparatively less attractive. The reason is that the showcase prizes are taxed on the basis of their (relatively high, non-discounted) retail prices.

[19]The effect of discounting the showcase value is equivalent to the effect of lowering contestants' perceived chance of winning the showcase after winning the SCSD. The present robustness test therefore also captures the possibility that this subjective probability is smaller than the 50 percent that we assumed in the main analyses.

[20]Extending the limited foresight model by allowing for omission bias now does yield a significant increase in explanatory power (LR test: $\chi^2(1) = 20.13$ and $p < 0.001$).

#### 4.4.3 Ignoring Bonus Prizes

A possible explanation for underspinning is that contestants are overly focused on reaching the final of the episode, and attach a relatively low weight to the possibility of winning one or two bonus prizes by obtaining a score of exactly 100. In this section, we re-estimate the structural models under the extreme assumption that contestants completely ignore the existence of the bonus prizes.[21]

Table 3, Panel C presents the results. As expected, ignoring the bonus prizes improves the overall fit of the baseline model. The fit of the models with omission bias, limited foresight, and the combination of these, however, is somewhat worse. More importantly, the limited foresight model still explains choices substantially better than the omission bias model (Vuong test: $Z = 9.08$, $p < 0.001$).[22]

#### 4.4.4 Other Robustness Checks

We perform four additional analyses to examine the robustness of our results to alternative modeling choices. First, instead of weighting the observations for Contestants 1 and 2 to correct for the imbalance in the sample sizes, we now give each observation equal weight. Second, we increase the flexibility of the structural models by allowing Contestant 1 and Contestant 2 to have different rationality parameters. As a third robustness check, we use the original, nominal monetary values instead of the inflation-corrected, real monetary values. Last, we assume that the expected showcase value equals the average stated retail price of all showcases in the previous season instead of the running season.

The four sets of results are in Table 4. In all cases, the limited foresight model provides a much better account of contestants' choices than the omission bias model (Vuong tests: all $Z > 2.54$, all $p < 0.006$).[23] The likelihood that a contestant myopically only considers the next stage of the game is barely affected by the alternative approaches: the limited foresight parameter is always close to the 41 percent that

[21] This robustness test also captures the possibility that contestants expect to derive relatively much utility from playing the *Showcase* final, for example because they are overconfident about their chances of winning the showcase, or because of the enjoyment associated with "winning the episode".

[22] Combining omission bias and limited foresight yields significantly more explanatory power than limited foresight alone (LR test: $\chi^2(1) = 4.50$ and $p = 0.034$). Note, however, that the estimated omission bias parameter is negative, which implies that people would have a preference for harmful actions over harmful inactions, and thus goes against the hypothesis.

[23] In three cases, adding the possibility of omission bias to the limited foresight model improves the empirical fit: under equal weighting, with separate rationality parameters, and when using nominal monetary values (LR tests: all $\chi^2(1) > 4.35$, all $p < 0.037$). There is no significant improvement when the expected showcase value is based on the previous season (LR test: $\chi^2(1) = 1.29$, $p = 0.255$).

Table 4: Estimation results under alternative modeling choices (2/2)

| | SPNE | | Baseline | | Omission bias | | Limited foresight | | OB & LF | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A: No weighting** | | | | | | | | | | |
| $\lambda$ | - | - | 1.423 | (0.027) | 1.677 | (0.035) | 1.764 | (0.047) | 1.776 | (0.045) |
| $\gamma$ | - | - | - | - | 1.051 | (0.046) | - | - | 0.385 | (0.113) |
| $\beta$ | - | - | - | - | - | - | 0.412 | (0.020) | 0.284 | (0.042) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -2,049 | | -1,814 | | -1,791 | | -1,786 | |
| AIC | - | | 4,101 | | 3,633 | | 3,586 | | 3,578 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.959 | 0.934 | 0.959 | 0.953 | 0.943 | 0.953 | 0.959 | 0.953 | 0.943 |
| Hit rate (difficult) | 0.729 | 0.795 | 0.729 | 0.795 | 0.817 | 0.705 | 0.817 | 0.795 | 0.817 | 0.703 |
| Brier score | 0.066 | 0.041 | 0.046 | 0.035 | 0.035 | 0.039 | 0.034 | 0.035 | 0.034 | 0.036 |
| Brier score (difficult) | 0.271 | 0.205 | 0.177 | 0.173 | 0.125 | 0.197 | 0.121 | 0.175 | 0.120 | 0.179 |
| Spinning bias | -0.063 | -0.014 | -0.041 | 0.001 | -0.008 | 0.027 | 0.005 | 0.001 | 0.003 | 0.011 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.217 | 0.010 | -0.066 | 0.147 | 0.011 | 0.017 | -0.003 | 0.069 |
| **Panel B: Separate rationality parameters** | | | | | | | | | | |
| $\lambda_1$ | - | - | 1.481 | (0.038) | 1.980 | (0.063) | 2.303 | (0.111) | 2.324 | (0.108) |
| $\lambda_2$ | - | - | 1.220 | (0.042) | 1.087 | (0.037) | 1.232 | (0.042) | 1.198 | (0.042) |
| $\gamma$ | - | - | - | - | 1.086 | (0.052) | - | - | 0.402 | (0.098) |
| $\beta$ | - | - | - | - | - | - | 0.424 | (0.021) | 0.303 | (0.036) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -1,954 | | -1,767 | | -1,731 | | -1,723 | |
| AIC | - | | 3,911 | | 3,539 | | 3,468 | | 3,454 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.959 | 0.934 | 0.959 | 0.953 | 0.945 | 0.953 | 0.959 | 0.953 | 0.943 |
| Hit rate (difficult) | 0.729 | 0.795 | 0.729 | 0.795 | 0.817 | 0.713 | 0.817 | 0.795 | 0.817 | 0.703 |
| Brier score | 0.066 | 0.041 | 0.046 | 0.035 | 0.035 | 0.038 | 0.035 | 0.035 | 0.035 | 0.035 |
| Brier score (difficult) | 0.271 | 0.205 | 0.178 | 0.173 | 0.125 | 0.186 | 0.125 | 0.173 | 0.124 | 0.175 |
| Spinning bias | -0.063 | -0.014 | -0.043 | 0.001 | -0.011 | 0.027 | 0.003 | 0.001 | 0.001 | 0.011 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.217 | 0.004 | -0.056 | 0.114 | 0.029 | 0.004 | 0.021 | 0.050 |
| **Panel C: Nominal monetary values** | | | | | | | | | | |
| $\lambda$ | - | - | 1.775 | (0.037) | 1.925 | (0.042) | 2.051 | (0.052) | 2.057 | (0.051) |
| $\gamma$ | - | - | - | - | 0.620 | (0.039) | - | - | 0.129 | (0.062) |
| $\beta$ | - | - | - | - | - | - | 0.405 | (0.024) | 0.351 | (0.035) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -2,013 | | -1,892 | | -1,837 | | -1,835 | |
| AIC | - | | 4,028 | | 3,788 | | 3,679 | | 3,676 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.959 | 0.934 | 0.959 | 0.950 | 0.943 | 0.953 | 0.959 | 0.953 | 0.943 |
| Hit rate (difficult) | 0.729 | 0.795 | 0.729 | 0.795 | 0.803 | 0.703 | 0.817 | 0.795 | 0.817 | 0.703 |
| Brier score | 0.066 | 0.041 | 0.046 | 0.036 | 0.037 | 0.038 | 0.035 | 0.036 | 0.034 | 0.036 |
| Brier score (difficult) | 0.271 | 0.205 | 0.175 | 0.175 | 0.130 | 0.186 | 0.121 | 0.176 | 0.120 | 0.177 |
| Spinning bias | -0.063 | -0.014 | -0.038 | 0.002 | -0.011 | 0.022 | 0.008 | 0.001 | 0.007 | 0.006 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.220 | 0.008 | -0.110 | 0.108 | 0.000 | 0.013 | -0.006 | 0.034 |
| **Panel D: Last season's showcase values** | | | | | | | | | | |
| $\lambda$ | - | - | 1.367 | (0.028) | 1.490 | (0.032) | 1.561 | (0.039) | 1.565 | (0.039) |
| $\gamma$ | - | - | - | - | 0.807 | (0.051) | - | - | 0.101 | (0.089) |
| $\beta$ | - | - | - | - | - | - | 0.409 | (0.024) | 0.374 | (0.039) |
| N | - | | 14,475 | | 14,475 | | 14,475 | | 14,475 | |
| Log-likelihood | - | | -1,959 | | -1,840 | | -1,789 | | -1,788 | |
| AIC | - | | 3,919 | | 3,683 | | 3,582 | | 3,583 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.935 | 0.959 | 0.935 | 0.959 | 0.952 | 0.943 | 0.952 | 0.959 | 0.952 | 0.948 |
| Hit rate (difficult) | 0.731 | 0.794 | 0.731 | 0.794 | 0.817 | 0.701 | 0.817 | 0.794 | 0.817 | 0.733 |
| Brier score | 0.065 | 0.041 | 0.046 | 0.035 | 0.037 | 0.037 | 0.034 | 0.035 | 0.034 | 0.035 |
| Brier score (difficult) | 0.269 | 0.206 | 0.176 | 0.173 | 0.132 | 0.186 | 0.120 | 0.174 | 0.120 | 0.174 |
| Spinning bias | -0.063 | -0.015 | -0.040 | 0.001 | -0.013 | 0.021 | 0.007 | 0.001 | 0.007 | 0.004 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.217 | 0.010 | -0.104 | 0.112 | 0.003 | 0.014 | -0.001 | 0.028 |

*Notes:* The table shows the results for four alternative modeling choices. Panel A shows the results when all individual observations for Contestants 1 and 2 are equally weighted, that is, without correcting for the imbalance in the sample sizes. Panel B shows the results when the rationality parameter, $\lambda$, is allowed to differ between Contestant 1 and Contestant 2. Panel C shows the results when nominal instead of real monetary values are used. Panel D shows the results under the assumption that the expected showcase value equals the average stated retail price of all showcases in the previous season instead of the running season. Other definitions are as in Table 2.

we found previously.

Finally, a possible concern may be that the order in which contestants take turns spinning the wheel is not random, but determined by the sum of the prizes they won in the previous games. This can be problematic if there is a relationship between contestants' prior winnings and their rationality. Such a relationship, however, is not very likely because the nature of the prior games is such that winnings are largely driven by luck. Moreover, empirically there is no evidence of such a relationship. When we regress the likelihood of following the optimal strategy on prior winnings, the regression coefficient is economically and statistically insignificant, regardless of whether we consider a linear or a log-linear relationship, and regardless of whether we consider all choices or difficult choices only (see Table A4 in the Appendix).

## 4.5 Learning

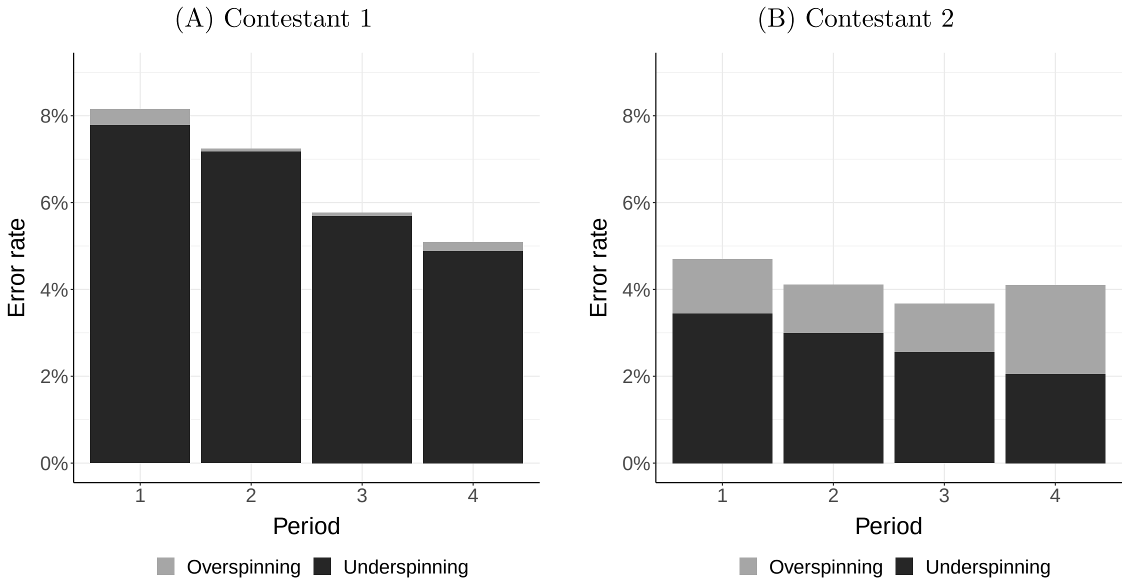
The SCSD has been running uninterruptedly for more than 40 years. This long history opens up the possibility to investigate whether behavior converges towards the rational equilibrium strategies over time, as contestants can learn about the game and the behavior of their opponents. In laboratory experiments, game theory tends to describe the behavior of experienced subjects better than that of inexperienced subjects (Fudenberg and Levine, 1998, 2009, 2016). Although SCSD contestants cannot gain experience themselves, they can potentially learn by observing the choices and outcomes of others (Duffy and Feltovich, 1999; Armantier, 2004; Simonsohn et al., 2008). Over time, the number of existing episodes has grown, and episodes have become more readily available. In addition, with the advent of the internet and modern communication technologies, people have become better able to share and discuss the optimal strategies.

To explore whether there is any evidence of learning, we divide our data into four different time periods: (i) seasons 1979-80 to 1992-93, (ii) 1993-94 to 2007-08, (iii) 2008-09 to 2014-15, and (iv) 2015-16 to 2020-21.[24]

For each of the four time periods, Figure 6 shows how often Contestants 1 and 2 deviate from the SPNE. For Contestant 1, there is a clear downward trend in the frequency of mistakes: the error rate decreases monotonically from 8.2 percent in Period 1 to 5.1 percent in Period 4. Nearly all of Contestant 1's deviations from the SPNE are underspinning errors, and the improvement over time almost fully reflects a reduction in underspinning. For Contestant 2, there is no clear trend in the overall quality of spinning decisions, with a constant error rate of roughly 4 percent

[24]We first separate the data for the two different bonus schemes, and then split the data for each bonus scheme into two periods of roughly equal length.

Figure 6: Deviations from the SPNE per period




*Notes:* The figure shows the fraction of spinning decisions by Contestant 1 (Panel A) and Contestant 2 (Panel B) that deviate from the SPNE, for four different time periods. The first period covers seasons 1979-80 to 1992-93, the second 1993-94 to 2007-08, the third 2008-09 to 2014-15, and the fourth 2015-16 to 2020-21. The dark gray parts of the bars reflect underspinning errors, the light gray parts reflect overspinning errors.

across all periods. In Periods 1 to 3 Contestant 2 exhibits more underspinning than overspinning, while in Period 4 these errors roughly balance out.

Of course comparing behavior in different time periods in this way is rather crude, because the costs of mistakes can be very different at different points in time due to the changing expected showcase value and the two different bonus schemes. The structural models account for such changes and confirm the improvement of Contestant 1's decisions. Table 5 shows the period-by-period estimation results for the limited foresight model. In line with the strong reduction of Contestant 1's underspinning relative to the SPNE, the fraction of spinning decisions that are made in accordance with limited foresight diminishes significantly over time: $\beta$ decreases monotonically from 65.4 to 18.3 percent.

The improved decision making over time is in line with learning. The results for the last period, however, show that even after several decades of *The Price Is Right*, a sizable proportion of contestants remain unable to follow the optimal strategies deriving from backward induction.

Table 5: Estimation results per period

| | Period 1 | | Period 2 | | Period 3 | | Period 4 | |
|---|---|---|---|---|---|---|---|---|
| $\lambda$ | 1.480 | (0.098) | 1.477 | (0.058) | 1.722 | (0.088) | 1.683 | (0.090) |
| $\beta$ | 0.654 | (0.067) | 0.491 | (0.039) | 0.345 | (0.045) | 0.183 | (0.050) |
| N | 2,012 | | 5,876 | | 3,770 | | 2,901 | |
| Log-likelihood | -260 | | -726 | | -425 | | -361 | |
| AIC | 525 | | 1,456 | | 854 | | 726 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.950 | 0.953 | 0.950 | 0.959 | 0.955 | 0.963 | 0.960 | 0.959 |
| Hit rate (difficult) | 0.838 | 0.781 | 0.807 | 0.786 | 0.831 | 0.798 | 0.826 | 0.815 |
| Brier score | 0.038 | 0.036 | 0.034 | 0.034 | 0.032 | 0.031 | 0.030 | 0.042 |
| Brier score (difficult) | 0.114 | 0.158 | 0.117 | 0.174 | 0.112 | 0.167 | 0.122 | 0.190 |
| Spinning bias | 0.012 | -0.006 | 0.011 | -0.003 | 0.003 | -0.001 | -0.002 | 0.019 |
| Spinning bias (difficult) | 0.019 | -0.021 | 0.006 | -0.015 | 0.000 | -0.006 | -0.041 | 0.105 |

*Notes:* The table shows the estimated parameters and the goodness-of-fit statistics of the structural model with limited foresight for four different time periods. The first period covers seasons 1979-80 to 1992-93, the second 1993-94 to 2007-08, the third 2008-09 to 2014-15, and the fourth 2015-16 to 2020-21. Other definitions are as in Table 2.

# 5 Conclusion and Discussion

The present paper examines high-stakes strategic decision making in the *Showcase Showdown* (SCSD), a sequential game of perfect information that is part of the long-running American TV game show *The Price Is Right*. The optimal strategies for this game can be found through backward induction. Most tests of the descriptive validity of backward induction as a solution concept rely on controlled laboratory experiments.[25] The SCSD provides an appealing alternative test bed, allowing for assessing the descriptive validity under conditions that are markedly different. The high stakes and ample learning opportunities provide a particularly benign setting for game-theoretic predictions to hold.

In spite of this, we find that contestants systematically deviate from the unique subgame perfect Nash equilibrium (SPNE). Their behavior is well explained by a modified agent quantal response equilibrium model that not only allows for random evaluation errors but also for limited foresight.[26] The results suggest that contestants are likely to simplify the decision problem by adopting a myopic representation and optimize their chances of beating the next contestant only. Our findings are robust to a rich variety of alternative modeling choices, including those regarding risk aversion. Omission bias plays little to no role, depending on the exact specification. In line with learning, we find that the frequency of deviations from the SPNE and the estimated degree of limited foresight decrease over the course of our sample period.

[25] One exception is Spenkuch et al. (2018), who find that voting behavior of US Senators during roll-call votes is largely consistent with the equilibrium predictions of a model in which the senators rely on backward induction.

[26] Chakraborty and Kendall (2023) analyze a single-player decision problem that requires subjects to reason contingently about their own decisions at hypothetical future events, and similarly find that behavior is best described by a model that combines QRE-like noise and limited foresight.

However, both systematic and non-systematic deviations remain commonplace, even after several decades.

Various published papers have derived the equilibrium strategies for the SCSD. Apparently, many contestants do not take heed of this information before coming on the show. This is consistent with research that demonstrates that people frequently do not use important and readily available information (for an overview, see Handel and Schwartzstein, 2018). Such ignorance is rational if the search costs outweigh the expected benefits (Stigler, 1961). For the SCSD, the expected benefits of thorough preparation are low: only six out of the several hundred audience members who travel to the recording studio actually play the SCSD, and only a fraction of those six end up in a relatively difficult choice situation where knowing the optimal strategy may truly be helpful. For many laypeople, the low ex-ante expected benefits probably do not outweigh the costs of looking up and reading a rather complicated academic paper.

The pattern of underspinning in the SCSD can be adequately captured by limited foresight. This well-documented bias, however, is not the only possible explanation. One alternative is rationality neglect. Several studies show that people tend to underestimate the rationality of their competition in strategic settings (Camerer and Lovallo, 1999; Weizsäcker, 2003; Rogers et al., 2009; Greenwood and Hanson, 2015). In the SCSD, overestimating the likelihood that others make mistakes generally lowers the perceived incentive to spin, especially for Contestant 1.

Underspinning might also result from overestimation of conjunctive events, which is the tendency to overestimate the likelihood of an event that requires the simultaneous occurrence of multiple conditions (Slovic, 1969; Cohen et al., 1972; Bar-Hillel, 1973; Vieider, 2011; Baillon et al., 2013). For Contestant 1, winning the SCSD requires the conjunction of beating Contestant 2 and beating Contestant 3. Overestimating conjunctive events thus elevates the subjective probability of winning, both for stopping and for spinning, but the impact is larger for stopping.[27]

In the present paper, we chose not to pursue these alternative explanations. Implementing them into our structural models is not straightforward and requires several additional modeling choices, thereby increasing both the degrees of freedom and the complexity. Moreover, the empirical fit of the limited foresight model leaves little room for improvement. Further research can use experiments with modifica-

[27]If Contestant 1 stops, the probability of winning equals the probability of the conjunctive event of beating both Contestant 2 and Contestant 3. If Contestant 1 spins, the probability of winning is an average of not only such conjunctive probabilities (for every potential score of 100 or less) but also the probability of zero (for every potential score exceeding the maximum of 100). Moreover, the conjunctive probabilities after spinning and improving the score are higher than those after stopping, limiting the scope for overestimation.

tions of the SCSD to try to shed further light on the possible explanations behind the deviations from the SPNE. It will, however, always remain unclear whether such laboratory findings generalize to our high-stakes setting, as the size and nature of deviations from the SPNE may well depend on the stakes.

*The Price Is Right* can be seen as an atypical setting to test the descriptive validity of backward induction, and critics may therefore view it as a negative distraction. However, novel settings should not be too easily dismissed as they can provide rare opportunities for relevant tests of economic theory (List, 2023). The SCSD uniquely allows for a large-scale analysis of strategic decision making at stakes that are impossible to replicate in the lab.

At the same time, following List (2023), it is important to explicitly consider how selection procedures and the naturalness of our setting may affect the generalizability of our results. Before contestants play the SCSD, they self-selected into the audience, were selected from the audience by the producers, and won a *One Bid* game. Unfortunately, it is unclear whether these elements of selection have led to any under- or overrepresentation of strategically sophisticated contestants. Selection effects, however, are inevitable in any lab or field setting. Moreover, SCSD contestants are quite diverse in terms of demographic characteristics, such as age, gender, ethnicity, and education, and as a group they seem to resemble a cross-section of the general population more closely than the subject pools of most laboratory experiments.

The setting in which contestants make their decisions—with a lively studio audience and camera's reminding them of the millions of TV viewers—likely induces stress. Psychological research indicates that the mere presence of others can facilitate performance in simple tasks but impair it in more complex ones (Zajonc, 1965; Bond and Titus, 1983). We cannot fully dismiss the impact the setting may have had on contestants, but prior research suggests that our findings are unlikely to be an artifact of the setting. Tenorio and Cason (2002) compare the behavior of laboratory subjects who play the SCSD to that of real contestants, and Antonovics et al. (2009), Healy and Noussair (2004), and Baltussen et al. (2016) make such a comparison for other games or game shows. None of these studies find that the patterns of behavior are different between the two types of settings. Moreover, every setting—including the experimental laboratory—is in some way special. It is impossible to study behavior under each and every possible set of conditions, and hence the optimal approach is to investigate if similar patterns are found in settings that are markedly different.

The finding that contestants in the SCSD often deviate from the optimal strategy and instead behave as if they adopt a simplified representation of the game adds to an ongoing debate about whether cognitive biases disappear in high-stake situations (Levitt and List, 2007*a*,*b*). Experimental research by Smith and Walker (1993), Cooper et al. (1999), Rapoport et al. (2003), and Parravano and Poulsen (2015) finds that the decisions of subjects tend to be closer to equilibrium play when the monetary incentives are higher. At the same time, Camerer and Hogarth (1999) and Enke et al. (2023) find that cognitive errors in experiments are largely impervious to the size of the stakes. Our results align with those of the latter two studies, and show that random and systematic violations of game-theoretic predictions abound in a high-stakes game that subjects can be expected to be highly familiar with.

# References


Anderson, S. P., Goeree, J. K. and Holt, C. A. (2001), 'Minimum-effort coordination games: Stochastic potential and logit equilibrium', *Games and Economic Behavior* **34**(2), 177–199.

Antonovics, K., Arcidiacono, P. and Walsh, R. (2009), 'The effects of gender interactions in the lab and in the field', *Review of Economics and Statistics* **91**(1), 152–162.

Armantier, O. (2004), 'Does observation influence learning?', *Games and Economic Behavior* **46**(2), 221–239.

Asch, D. A., Baron, J., Hershey, J. C., Kunreuther, H., Meszaros, J., Ritov, I. and Spranca, M. (1994), 'Omission bias and pertussis vaccination', *Medical Decision Making* (14), 118–123.

Atanasov, P. D., Dana, J. D. and Klein Teeselink, B. (2023), 'Own-gender favoritism in high stakes decisions: Taste-based discrimination on The Price Is Right'. *Economic Journal*, forthcoming.

Baillon, A., Selim, A. and van Dolder, D. (2013), 'On the social nature of eyes: The effect of social cues in interaction and individual choice tasks', *Evolution and Human Behavior* **34**(2), 146–154.

Baltussen, G., van den Assem, M. J. and van Dolder, D. (2016), 'Risky choice in the limelight', *Review of Economics and Statistics* **98**(2), 318–332.

Bar-Hillel, M. (1973), 'On the subjective probability of compound events', *Organizational Behavior and Human Performance* **9**(3), 396–406.

Baranski, A. and Reuben, E. (2023), Competing for proposal rights: Theory and experimental evidence. Working paper.

Barreda-Tarrazona, I., Kundu, T. and Østbye, S. (2021), 'On rational forward-looking behavior in economic geography: An experimental analysis', *Regional Science and Urban Economics* **87**, 103654.

Belot, M., Bhaskar, V. and van de Ven, J. (2010), 'Promises and cooperation: Evidence from a TV game show', *Journal of Economic Behavior and Organization* **73**(3), 396–405.

Bennett, R. W. and Hickman, K. A. (1993), 'Rationality and The Price Is Right', *Journal of Economic Behavior and Organization* **21**(1), 99–105.

Berk, J. B., Hughson, E. and Vandezande, K. (1996), 'The price is right, but are the bids? An investigation of rational decision theory', *American Economic Review* **86**(4), 954–970.

Binmore, K. (1999), 'Why experiment in economics?', *Economic Journal* **109**(453), 16–24.

Binmore, K., McCarthy, J., Ponti, G., Samuelson, L. and Shaked, A. (2002), 'A backward induction experiment', *Journal of Economic Theory* **104**(1), 48–88.

Bombardini, M. and Trebbi, F. (2012), 'Risk aversion and expected utility theory: An experiment with large and small stakes', *Journal of the European Economic Association* **10**(6), 1348–1399.

Bond, C. F. and Titus, L. J. (1983), 'Social facilitation: A meta-analysis of 241 studies', *Psychological Bulletin* **94**(2), 265–292.

Bossaerts, P. L., Fattinger, F., van den Bogaerde, F. and Yang, W. (2022), Asset pricing in a world of imperfect foresight. Working paper.

Brier, G. W. (1950), 'Verification of forecasts expressed in terms of probability', *Monthly Weather Review* (75), 1–3.

Burks, R. E. and Jaye, M. J. (2012), 'The price is right again', *Journal of Statistics Education* **20**(2).

Buser, T., van den Assem, M. J. and van Dolder, D. (2023), 'Gender and willingness to compete for high stakes', *Journal of Economic Behavior and Organization* **206**, 350–370.

Cai, H. and Wang, J. T. Y. (2006), 'Overcommunication in strategic information transmission games', *Games and Economic Behavior* **56**(1), 7–36.

Camerer, C. F. (2015), The promise and success of lab–field generalizability in experimental economics: A critical reply to Levitt and List, *in* G. R. Fréchette and A. Schotter, eds, 'Handbook of Experimental Economic Methodology', Oxford University Press.

Camerer, C. F. and Hogarth, R. M. (1999), 'The effects of financial incentives in experiments: A review and capital-labor-production framework', *Journal of Risk and Uncertainty* **19**, 7–42.

Camerer, C. and Lovallo, D. (1999), 'Overconfidence and excess entry: An experimental approach', *American Economic Review* **89**(1), 306–318.

Capra, C. M., Goeree, J. K., Gomez, R. and Holt, C. A. (1999), 'Anomalous behavior in a traveler's dilemma?', *American Economic Review* **89**(3), 678–690.

Carlin, B. I. and Robinson, D. T. (2009), 'Fear and loathing in Las Vegas: Evidence from blackjack tables', *Judgment and Decision Making* **4**(5), 385–396.

Cason, T. N. and Reynolds, S. S. (2005), 'Bounded rationality in laboratory bargaining with asymmetric information', *Economic Theory* **25**(3), 553–574.

Chakraborty, A. and Kendall, C. (2023), Noisy foresight. Working paper.

Chen, H. C., Friedman, J. W. and Thisse, J. F. (1997), 'Boundedly rational Nash equilibrium: A probabilistic choice approach', *Games and Economic Behavior* **18**(1), 32–54.

Coe, P. R. and Butterworth, W. (1995), 'Optimal stopping in the Showcase Showdown', *American Statistician* **49**(3), 271–275.

Cohen, J., Chesnick, E. and Haran, D. (1972), 'A confirmation of the inertial-Ψ effect in sequential choice and decision', *British Journal of Psychology* **63**(1), 41–46.

Cooper, D. J., Kagel, J. H.and Lo, W. and Gu, Q. L. (1999), 'Gaming against managers in incentive systems: Experimental results with Chinese students and Chinese managers', *American Economic Review* **89**(4), 781–804.

Deck, C. A. (2001), 'A test of game-theoretic and behavioral models of play in exchange and insurance environments', *American Economic Review* **91**(5), 1546–1555.

DiBonaventura, M. D. and Chapman, G. B. (2008), 'Do decision biases predict bad decisions? Omission bias, naturalness bias, and influenza vaccination', *Medical Decision Making* **28**(4), 532–539.

Dixit, A. (1982), 'Recent developments in oligopoly theory', *American Economic Review* **72**(2), 12–17.

Duffy, J. and Feltovich, N. (1999), 'Does observation of others affect learning in strategic environments? An experimental study', *International Journal of Game Theory* **28**(1), 131–152.

Dufwenberg, M. and Van Essen, M. (2018), 'King of the hill: Giving backward induction its best shot', *Games and Economic Behavior* **112**, 125–138.

Eberhardt, I., Smeets, P., van den Assem, M. J. and van Dolder, D. (2024), Impact or responsibility? Giving in a televised natural experiment. Working paper.

Enke, B., Gneezy, U., Hall, B., Martin, D. C., Nelidov, V., Offerman, T. and van de Ven, J. (2023), 'Cognitive biases: Mistakes or missing stakes?', *Review of Economics and Statistics* **105**(4), 818–832.

Falk, A. and Heckman, J. J. (2009), 'Lab experiments are a major source of knowledge in the social sciences', *Science* **326**(5952), 535–538.

Fehr, E., Powell, M. and Wilkening, T. (2021), 'Behavioral constraints on the design of subgame-perfect implementation mechanisms', *American Economic Review* **111**(4), 1055–91.

Feldman, G., Kutscher, L. and Yay, T. (2020), 'Omission and commission in judgment and decision making: Understanding and linking action-inaction effects using the concept of normality', *Social and Personality Psychology Compass* **14**(8), e12557.

Fey, M., McKelvey, R. D. and Palfrey, T. R. (1996), 'An experimental study of constant-sum centipede games', *International Journal of Game Theory* **25**(3), 269–287.

Fudenberg, D. and Levine, D. K. (1998), *The Theory of Learning in Games*, MIT Press.

Fudenberg, D. and Levine, D. K. (2009), 'Learning and equilibrium', *Annual Review of Economics* **1**(1), 385–420.

Fudenberg, D. and Levine, D. K. (2016), 'Whither game theory? Towards a theory of learning in games', *Journal of Economic Perspectives* **30**(4), 151–70.

Gabaix, X. and Laibson, D. (2005), Bounded rationality and directed cognition. Working paper.

Gabaix, X., Laibson, D., Moloche, G. and Weinberg, S. (2006), 'Costly information acquisition: Experimental analysis of a boundedly rational model', *American Economic Review* **96**(4), 1043–1068.

Gertner, R. (1993), 'Game shows and economic behavior: Risk-taking on Card Sharks', *Quarterly Journal of Economics* **108**(2), 507–521.

Goeree, J. K. and Holt, C. A. (2001), 'Ten little treasures of game theory and ten intuitive contradictions', *American Economic Review* **91**(5), 1402–1422.

Goeree, J. K., Holt, C. A. and Palfrey, T. R. (2002), 'Quantal response equilibrium and overbidding in private-value auctions', *Journal of Economic Theory* **104**(1), 247–272.

Goeree, J. K., Holt, C. A. and Palfrey, T. R. (2003), 'Risk averse behavior in generalized matching pennies games', *Games and Economic Behavior* **45**(1), 97–113.

Goeree, J. K., Holt, C. A. and Palfrey, T. R. (2005), 'Regular quantal response equilibrium', *Experimental Economics* **8**(4), 347–367.

Goeree, J. K., Holt, C. A. and Palfrey, T. R. (2016), *Quantal Response Equilibrium*, Princeton University Press.

Goeree, J. K., Holt, C. A. and Palfrey, T. R. (2020), Stochastic game theory for social science: A primer on quantal response equilibrium, *in* C. M. Capra, R. T. Croson, M. L. Rigdon and T. S. Rosenblat, eds, 'Handbook of Experimental Game Theory', Edward Elgar Publishing.

Goeree, J. K., Holt, C. A. and Smith, A. M. (2017), 'An experimental examination of the volunteer's dilemma', *Games and Economic Behavior* **102**, 303–315.

Greenwood, R. and Hanson, S. G. (2015), 'Waves in ship prices and investment', *Quarterly Journal of Economics* **130**(1), 55–110.

Grosjean, J. H. (1998), 'Beating the Showcase Showdown', *Chance* **11**(1), 14–19.

Haile, P. A., Hortaçsu, A. and Kosenok, G. (2008), 'On the empirical content of quantal response equilibrium', *American Economic Review* **98**(1), 180–200.

Hallsworth, M., List, J. A., Metcalfe, R. D. and Vlaev, I. (2023), The making of homo honoratus: From omission to commission. *Journal of Consumer Psychology*, forthcoming.

Handel, B. and Schwartzstein, J. (2018), 'Frictions or mental gaps: What's behind the information we (don't) use and when do we care?', *Journal of Economic Perspectives* **32**(1), 155–178.

Healy, P. and Noussair, C. (2004), 'Bidding behavior in the Price Is Right game: An experimental study', *Journal of Economic Behavior and Organization* **54**(2), 231–247.

Hogarth, R. M., Karelaia, N. and Trujillo, C. A. (2012), 'When should I quit? Gender differences in exiting competitions', *Journal of Economic Behavior and Organization* **83**(1), 136–150.

Jackson, M. O. and Wolinsky, A. (1996), 'A strategic model of social and economic networks', *Journal of Economic Theory* **71**(1), 44–74.

Jehiel, P. (1995), 'Limited horizon forecast in repeated alternate games', *Journal of Economic Theory* **67**(2), 497–519.

Jehiel, P. (1998), 'Learning to play limited forecast equilibria', *Games and Economic Behavior* **22**(2), 274–298.

Jehiel, P. (2001), 'Limited foresight may force cooperation', *Review of Economic Studies* **68**(2), 369–391.

Johnson, E. J., Camerer, C. F., Sen, S. and Rymon, T. (2002), 'Detecting failures of backward induction: Monitoring information search in sequential bargaining', *Journal of Economic Theory* **104**(1), 16–47.

Ke, S. (2019), 'Boundedly rational backward induction', *Theoretical Economics* **14**(1), 103–134.

Levitt, S. D. (2004), 'Testing theories of discrimination: Evidence from Weakest Link', *Journal of Law and Economics* **47**(2), 431–453.

Levitt, S. D. and List, J. A. (2007*a*), 'On the generalizability of lab behaviour to the field', *Canadian Journal of Economics* **40**(2), 347–370.

Levitt, S. D. and List, J. A. (2007*b*), 'What do laboratory experiments measuring social preferences reveal about the real world?', *Journal of Economic Perspectives* **21**(2), 153–174.

Levitt, S. D., List, J. A. and Sadoff, S. E. (2011), 'Checkmate: Exploring backward induction among chess players', *American Economic Review* **101**(2), 975–990.

List, J. A. (2006), 'Friend or Foe? A natural experiment of the prisoner's dilemma', *Review of Economics and Statistics* **88**(3), 463–471.

List, J. A. (2023), Non est disputandum de generalizability? A glimpse into the external validity trial. Working paper.

Mantovani, M. (2016), Limited foresight in sequential games: An experiment. Working paper.

McKelvey, R. D. and Palfrey, T. R. (1992), 'An experimental study of the centipede game', *Econometrica* **60**(4), 803–836.

McKelvey, R. D. and Palfrey, T. R. (1995), 'Quantal response equilibria for normal form games', *Games and Economic Behavior* **10**(1), 6–38.

McKelvey, R. D. and Palfrey, T. R. (1998), 'Quantal response equilibria for extensive form games', *Experimental Economics* **1**(1), 9–41.

McKelvey, R. D. and Patty, J. W. (2006), 'A theory of voting in large elections', *Games and Economic Behavior* **57**(1), 155–180.

Metrick, A. (1995), 'A natural experiment in Jeopardy', *American Economic Review* **85**(1), 240–253.

Moinas, S. and Pouget, S. (2013), 'The bubble game: An experimental study of speculation', *Econometrica* **81**(4), 1507–1539.

Moskowitz, T. J. and Wertheim, L. J. (2011), *Scorecasting: The hidden influences behind how sports are played and games are won*, New York: Crown Pub.

Neelin, J., Sonnenschein, H. and Spiegel, M. (1988), 'A further test of noncooperative bargaining theory: Comment', *American Economic Review* **78**(4), 824–836.

Oberholzer-Gee, F., Waldfogel, J. and White, M. W. (2010), 'Friend or Foe? Cooperation and learning in high-stakes games', *Review of Economics and Statistics* **92**(1), 179–187.

OECD (2021), 'Main economic indicators - complete database', Accessed on 12 July 2021.

Parravano, M. and Poulsen, O. (2015), 'Stake size and the power of focal points in coordination games: Experimental evidence', *Games and Economic Behavior* **94**, 191–199.

Post, G. T., van den Assem, M. J., Baltussen, G. and Thaler, R. H. (2008), 'Deal or no deal? Decision making under risk in a large-payoff game show', *American Economic Review* **98**(1), 38–71.

Rampal, J. (2022), 'Limited foresight equilibrium', *Games and Economic Behavior* **132**, 166–188.

Rapoport, A., Stein, W. E., Parco, J. E. and Nicholas, T. E. (2003), 'Equilibrium play and adaptive learning in a three-person centipede game', *Games and Economic Behavior* **43**(2), 239–265.

Ritov, I. and Baron, J. (1990), 'Reluctance to vaccinate: Omission bias and ambiguity', *Journal of Behavioral Decision Making* **3**(4), 263–277.

Ritov, I. and Baron, J. (1992), 'Status-quo and omission biases', *Journal of Risk and Uncertainty* **5**(1), 49–61.

Rogers, B. W., Palfrey, T. R. and Camerer, C. F. (2009), 'Heterogeneous quantal response equilibrium and cognitive hierarchies', *Journal of Economic Theory* **144**(4), 1440–1467.

Rosenthal, R. W. (1981), 'Games of perfect information, predatory pricing and the chain-store paradox', *Journal of Economic Theory* **25**(1), 92–100.

Rubinstein, A. (1982), 'Perfect equilibrium in a bargaining model', *Econometrica* **50**(1), 97–109.

Selten, R. (1978), 'The chain store paradox', *Theory and Decision* **2**(9), 127–159.

Simonsohn, U., Karlsson, N., Loewenstein, G. and Ariely, D. (2008), 'The tree of experience in the forest of information: Overweighing experienced relative to observed information', *Games and Economic Behavior* **62**(1), 263–286.

Slovic, P. (1969), 'Manipulating the attractiveness of a gamble without changing its expected value.', *Journal of Experimental Psychology* **79**(1), 139–145.

Smith, V. L. and Walker, J. M. (1993), 'Monetary rewards and decision cost in experimental economics', *Economic Inquiry* **31**(2), 245–261.

Spenkuch, J. L., Montagnes, B. P. and Magleby, D. B. (2018), 'Backward induction in the wild? Evidence from sequential voting in the US Senate', *American Economic Review* **108**(7), 1971–2013.

Spranca, M., Minsk, E. and Baron, J. (1991), 'Omission and commission in judgment and choice', *Journal of Experimental Social Psychology* **27**(1), 76–105.

Stigler, G. J. (1961), 'The economics of information', *Journal of Political Economy* **69**(3), 213–225.

Swenson, D. (2015), 'Optimal strategy in the Price Is Right Showcase Showdown: A module for students of calculus and probability', *PRIMUS* **25**(7), 578–595.

Tenorio, R. and Cason, T. N. (2002), 'To spin or not to spin? Natural and laboratory experiments from The Price Is Right', *Economic Journal* **112**(476), 170–195.

Turmunkh, U., van den Assem, M. J. and van Dolder, D. (2019), 'Malleable lies: Communication and cooperation in a high stakes TV game show', *Management Science* **65**(10), 4795–4812.

van den Assem, M. J., van Dolder, D. and Thaler, R. H. (2012), 'Split or Steal? Cooperative behavior when the stakes are large', *Management Science* **58**(1), 2–20.

van Dolder, D., van den Assem, M. J., Camerer, C. F. and Thaler, R. H. (2015), 'Standing united or falling divided? High stakes bargaining in a TV game show', *American Economic Review* **105**(5), 402–407.

Vieider, F. M. (2011), 'Separating real incentives and accountability', *Experimental Economics* **14**, 507–518.

von Stackelberg, H. (1934), *Marktform und Gleichgewicht*, Vienna: Springer.

Walker, J., Risen, J. L., Gilovich, T. and Thaler, R. (2018), 'Sudden death aversion: Avoiding superior options because they feel riskier', *Journal of Personality and Social Psychology* **115**(3), 363–378.

Weizsäcker, G. (2003), 'Ignoring the rationality of others: Evidence from experimental normal-form games', *Games and Economic Behavior* **44**(1), 145–171.

Zajonc, R. B. (1965), 'Social facilitation: A solution is suggested for an old unresolved social psychological problem', *Science* **149**(3681), 269–274.

# Appendix

Table A1: Data coverage per season

| Season | Episodes | SCSDs | Showcases |
|---|---|---|---|
| 1979-1980 | 32 | 61 | 60 |
| 1980-1981 | 30 | 54 | 58 |
| 1981-1982 | 31 | 59 | 60 |
| 1982-1983 | 192 | 368 | 378 |
| 1983-1984 | 93 | 174 | 179 |
| 1984-1985 | 16 | 28 | 31 |
| 1985-1986 | 15 | 27 | 29 |
| 1986-1987 | 65 | 122 | 118 |
| 1987-1988 | 22 | 41 | 39 |
| 1988-1989 | 25 | 49 | 49 |
| 1989-1990 | 25 | 49 | 49 |
| 1990-1991 | 15 | 27 | 28 |
| 1991-1992 | 83 | 163 | 163 |
| 1992-1993 | 80 | 151 | 160 |
| 1993-1994 | 75 | 134 | 92 |
| 1994-1995 | 111 | 207 | 175 |
| 1995-1996 | 112 | 201 | 223 |
| 1996-1997 | 140 | 227 | 268 |
| 1997-1998 | 116 | 189 | 219 |
| 1998-1999 | 130 | 240 | 256 |
| 1999-2000 | 134 | 264 | 265 |
| 2000-2001 | 171 | 309 | 341 |
| 2001-2002 | 182 | 363 | 364 |
| 2002-2003 | 173 | 345 | 346 |
| 2003-2004 | 170 | 340 | 294 |
| 2004-2005 | 159 | 316 | 255 |
| 2005-2006 | 168 | 336 | 267 |
| 2006-2007 | 150 | 249 | 125 |
| 2007-2008 | 179 | 354 | 353 |
| 2008-2009 | 190 | 372 | 375 |
| 2009-2010 | 188 | 375 | 374 |
| 2010-2011 | 189 | 370 | 372 |
| 2011-2012 | 192 | 376 | 382 |
| 2012-2013 | 186 | 361 | 370 |
| 2013-2014 | 193 | 378 | 386 |
| 2014-2015 | 187 | 369 | 373 |
| 2015-2016 | 193 | 383 | 385 |
| 2016-2017 | 178 | 355 | 356 |
| 2017-2018 | 175 | 350 | 350 |
| 2018-2019 | 176 | 350 | 351 |
| 2019-2020 | 158 | 313 | 316 |
| 2020-2021 | 136 | 272 | 272 |

*Notes:* The table displays the coverage of our sample per season. *Episodes* is the number of episodes for which we have the data for at least one of the two SCSDs. *SCSDs* is the number of SCSDs for which we have all spinning decisions and outcomes. *Showcases* is the number of showcases for which we know the stated retail price.

Table A2: Costs of deviations from the SPNE

| First spin | C1 | C2 |
|---|---|---|
| 5 | $3,264 | $5,223 |
| 10 | $3,247 | $5,123 |
| 15 | $3,216 | $4,939 |
| 20 | $3,166 | $4,668 |
| 25 | $3,094 | $4,306 |
| 30 | $2,995 | $3,850 |
| 35 | $2,861 | $3,295 |
| 40 | $2,684 | $2,638 |
| 45 | $2,454 | $1,874 |
| 50 | $2,161 | $1,002 |
| 55 | $1,800 | $55 |
| 60 | $1,189 | $1,008 |
| 65 | $393 | $2,189 |
| 70 | $555 | $3,493 |
| 75 | $1,796 | $4,923 |
| 80 | $3,333 | $6,484 |
| 85 | $5,211 | $8,179 |
| 90 | $7,482 | $10,012 |
| 95 | $10,199 | $11,987 |
| 100 | $13,585 | $14,269 |

*Notes:* The table shows the costs of deviating from the SPNE when the expected showcase value is $30,000 and bonus prizes are $1,000, $10,000, and $25,000. For Contestant 2 (C2), the costs are for choice situations where their first-spin outcome beats the score of Contestant 1 (C1). Under these conditions, the stopping thresholds are 70 (C1) and 60 (C2).

Table A3: Estimation results under different levels of risk aversion

| | SPNE | | Baseline | | Omission bias | | Limited foresight | | OB & LF | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel A: Medium risk aversion** | | | | | | | | | | |
| $\lambda$ | - | - | 5.141 | (0.103) | 5.675 | (0.121) | 5.943 | (0.149) | 5.971 | (0.149) |
| $\gamma$ | - | - | - | - | 0.235 | (0.013) | - | - | 0.046 | (0.024) |
| $\beta$ | - | - | - | - | - | - | 0.439 | (0.024) | 0.379 | (0.039) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -1,981 | | -1,836 | | -1,788 | | -1,786 | |
| AIC | - | | 3,965 | | 3,676 | | 3,579 | | 3,578 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.959 | 0.934 | 0.959 | 0.952 | 0.943 | 0.953 | 0.959 | 0.953 | 0.947 |
| Hit rate (difficult) | 0.729 | 0.795 | 0.729 | 0.795 | 0.815 | 0.703 | 0.817 | 0.795 | 0.817 | 0.728 |
| Brier score | 0.066 | 0.041 | 0.047 | 0.035 | 0.037 | 0.037 | 0.034 | 0.035 | 0.034 | 0.035 |
| Brier score (difficult) | 0.271 | 0.205 | 0.182 | 0.172 | 0.133 | 0.185 | 0.121 | 0.173 | 0.120 | 0.173 |
| Spinning bias | -0.063 | -0.014 | -0.043 | -0.001 | -0.014 | 0.021 | 0.008 | -0.001 | 0.007 | 0.004 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.229 | -0.001 | -0.103 | 0.113 | 0.008 | 0.004 | 0.001 | 0.027 |
| **Panel B: Low risk aversion** | | | | | | | | | | |
| $\lambda$ | - | - | 2.155 | (0.044) | 2.344 | (0.050) | 2.456 | (0.061) | 2.461 | (0.061) |
| $\gamma$ | - | - | - | - | 0.508 | (0.032) | - | - | 0.055 | (0.057) |
| $\beta$ | - | - | - | - | - | - | 0.405 | (0.024) | 0.375 | (0.039) |
| N | - | | 14,559 | | 14,559 | | 14,559 | | 14,559 | |
| Log-likelihood | - | | -1,956 | | -1,839 | | -1,790 | | -1,790 | |
| AIC | - | | 3,915 | | 3,682 | | 3,585 | | 3,586 | |
| | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 | C1 | C2 |
| Hit rate | 0.934 | 0.959 | 0.934 | 0.960 | 0.952 | 0.943 | 0.953 | 0.960 | 0.953 | 0.945 |
| Hit rate (difficult) | 0.729 | 0.795 | 0.729 | 0.796 | 0.815 | 0.701 | 0.817 | 0.796 | 0.817 | 0.714 |
| Brier score | 0.066 | 0.041 | 0.045 | 0.035 | 0.037 | 0.037 | 0.034 | 0.035 | 0.034 | 0.035 |
| Brier score (difficult) | 0.271 | 0.205 | 0.175 | 0.174 | 0.133 | 0.186 | 0.120 | 0.175 | 0.120 | 0.175 |
| Spinning bias | -0.063 | -0.014 | -0.040 | 0.002 | -0.014 | 0.021 | 0.007 | 0.002 | 0.006 | 0.004 |
| Spinning bias (difficult) | -0.258 | -0.066 | -0.216 | 0.011 | -0.104 | 0.112 | 0.003 | 0.015 | 0.000 | 0.027 |

*Notes:* The table shows the results for medium and low degrees of risk aversion. Panel A (Panel B) shows the results under the assumption that contestants have CARA utility, with a certainty equivalent of \$5,000 (\$10,000) for a 50-50 lottery of winning \$25,000 or \$0.

Table A4: Optimal choices and prior winnings

| | All choices | | Difficult choices | |
|---|---|---|---|---|
| | Model 1 | Model 2 | Model 3 | Model 4 |
| Prior winnings | 0.00003 | | 0.0003 | |
| | (0.00005) | | (0.0002) | |
| ln(Prior winnings) | | 0.001 | | 0.001 |
| | | (0.002) | | (0.007) |
| Fixed effects | Yes | Yes | Yes | Yes |
| Observations | 12,665 | 12,665 | 2,491 | 2,491 |

*Notes:* The table shows regression results for the relationship between the optimality of play and prior winnings. The dependent variable is a dummy variable that takes the value of 1 if the contestant follows the optimal strategy according to the SPNE, and 0 otherwise. *Prior winnings* is the inflation-corrected monetary value of the prizes won by the contestant prior to the SCSD, in thousands of dollars. *ln(Prior winnings)* is the natural logarithm of *Prior winnings*. Fixed effects allow for differences in the average likelihood of a departure from optimality across first-spin outcomes, separately for Contestant 1 and for Contestant 2, and, in the case of Contestant 2, for whether their first spin beats or ties the previous contestant's score. Models 1 and 2 are estimated on all observations for which prior winnings are available in our data; Models 3 and 4 are estimated on relatively difficult choice situations only. Difficult choices are choices where the first-spin outcome is no more than two steps below the stopping threshold and no more than one step above it. Standard errors are in parentheses.