German Economic Review 3(2): 223–245

Parimutuel Lotteries: Gamblers' Behavior and the Demand for Tickets

Ursula Hauser-Rethaller
Institute for Advanced Studies, Vienna

Ulrich König
University of Vienna

Abstract. *In a parimutuel lottery, players face a strategic situation. We investigate how rational lottery players should choose combinations of numbers. Using data from the Austrian Lotto we compare this to actual behavior. We propose a relationship between the number of tickets and the expected loss of taking part, based on both theoretical and empirical findings about players' behavior. Rollovers introduce exogenous price variation allowing to estimate properties of a demand function sensitive to the expected loss. Contrary to previous work our model accounts for conscious selection.*

1. INTRODUCTION

After decades of social outlawing, restrictive regulation or even prohibition, the gambling industry has experienced a tremendous increase in turnover and popularity in Europe as well as in the USA. This development may be due to a changing attitude within the population versus an activity promising a monetary return without effort. Even more importantly, lotteries are a perfect vehicle for governments' needs to finance rising public expenditures without introducing unpopular taxes.[1] We have thus witnessed the establishment of state-run or state-controlled agencies in most European countries,which offer a broad and continuously more differentiated spectrum of gambles.[2]

Large-scale state-run gambles are often organized as *parimutuel* gambles. Their main characteristic is that by definition, the revenue of the provider is a

1. The tax rate for the Austrian lotto is 38%, the tax rate for the Austrian scratchcard game is 27% of the purchasing price. This is comparable to other high-taxed products as tobacco (34–47%). The usefulness of lotteries for taxation is reflected in the old ironic saying:

   Lotto's a taxation
   On all fools in the nation
   But heaven be praised
   It's so easily raised.

2. For a survey of US lotteries see Clotfelter and Cook (1989). A description of the development and status quo of European lottery markets can be found in Boss *et al.* (1997).

fixed percentage of the total handle. From the provider's standpoint, this does not only mean more certain returns, it also secures that he has no interest in deliberately influencing the outcome of the draw, thereby improving his credibility amongst participating players. As the game design allows the number of winning tickets to vary, the distribution of the prize pool amongst winners will lead to variation in the size of prizes.

In a fixed-odds gamble with a draw depending purely on chance, the individual player cannot influence the expected percentage loss of taking part, this being equal to the take-out rate. For instance, when playing Roulette, it is completely irrelevant whether one chooses to bet on 14 or 27. Parimutuel gambles, however, are different. Even if the determination of winning tickets is completely random (we will call this type of gamble a lottery), the decision which numbers to predict potentially influences the expected payoff of the bet. While in a fixed-odds gamble participation is essentially a bet with the provider as a counterpart, the provider of a parimutuel gamble has just a passive role. The participants' decisions mutually influence each others' expected payoff, leaving them in a strategic situation. Similarities to financial markets are evident here. As, in contrast to stock markets, the uncertainty about future payoffs is resolved rather quickly and one does not need to worry about retrading effects, parimutuel betting markets have often served as a testing area for issues like weak and strong forms of market efficiency (see, for instance, Gabriel and Marsden, 1990, or Thaler and Ziemba, 1988). An 'efficient' gambling market can be imagined as one where the different outcomes of the draw (win of horse A, a combination of six numbers in Lotto etc.) are neither over- nor underbet by the participants. For a lottery, where all outcomes have the same objective probabilities, we will show that the situation where all players bet ('predict') combinations of numbers at random from an equal distribution, indeed constitutes a Nash equilibrium of the underlying coordination game, although there will be other equilibria that Pareto-dominate that one.

Another interesting aspect is the price of gambling. In our context, the wager should not be seen as the price. It should rather be interpreted as the minimal, indivisible unit of the good 'gambling'. The underlying model is one of consumers deriving utility from the process of gambling itself, in the sense of Conlisk (1993), rather than from the uncertain payoffs alone.[3] Disregarding their attitude towards risk, these consumers are prepared to sacrifice a certain 'entrance fee' to be able to participate in a (then fair) gamble with given characteristics. This entrance fee is equal to the expected loss per ticket and under suitable assumptions one would arrive at a market demand for tickets, sensitive to this expected loss (or price of participation).

Unfortunately, the take-out rate as the main characteristic determining the expected loss is not changed frequently enough in order to infer demand

3. Excitement, thrill and debates about lottery outcomes in the peer group are possible factors contributing to this intrinsic utility.

elasticities with any degree of statistical certainty (Suits, 1979). The resort to cross-country studies, on the other hand, faces the problem of simultaneous changes in game design and clientele (Cook and Clotfelter, 1993). There is, of course, no reason to believe that a lottery with lots of smaller prizes is viewed as a perfect substitute for a gamble with fewer, larger prizes. Moreover, the take-out rate does not exhibit much variation across different countries for a given type of gamble. National lotteries are typically run at a take-out rate of around 50%. Across different gambles, take-out rate differentials are substantial, but still gambles with high expected loss persist. Roulette has a low take-out rate (1.39–2.7%) whereas many instant lotteries have take-out rates of up to 60%. This underlines that gambles with different designs are not close substitutes.

However, parimutuel lotteries allow for a different approach. It is possible that there exists no winning ticket. As the take-out is fixed as a percentage of the handle, the undistributed prize pool must be paid out to the participants in a different way. Usually, this is effected by a *rollover*. So if there is no winner, the undistributed prize pool will be added to the next round's jackpot. We show that, due to this structure, the expected loss from buying a ticket depends on the number of participating tickets as well as on the size of the rollover from previous rounds. Rollovers will therefore introduce exogenous price variation, while other characteristics of the gamble remain the same. The data for the total number of purchased tickets then allows us to infer a demand relationship. While in rounds with no previous rollover, higher participation will lead to a lower expected loss per ticket the effect of increased participation in rounds with a previous rollover on the price of a ticket is ambiguous. Cook and Clotfelter (1993) have called this 'Peculiar Scale Economies of Lotto'. When estimating a demand curve using data from Massachusetts Lotto, they acknowledge that non-random selection of numbers by lotto players will bias their results but do not attempt to account for this phenomenon in their model. In two recent papers, Farrell and Walker (2000) and Walker (1998) discuss that problem. From the data of the UK Lottery – drawn numbers and final odds – they try to infer which numbers are more popular than others. Finding that non-random ('conscious') selection is prevalent, they revise the estimation of demand parameters.

A shortcoming of these studies is that data are usually limited to sizes of prizes in different rounds. Basically, the more variation one observes in these, the more uneven is coverage of the combinations by the players. Although our data set is longer than the one used by Farrell and Walker (570 observations versus 92) it is still tiny when compared to the huge number of possible combinations. Naturally, this allows for crude guesses about the true behavior of lotto players only.[4] We propose a simple alternative extension of the pricing

4. Simon (1998) had the rare opportunity to examine the full distribution of played combinations for a single round of the UK lottery. It turns out that it is still difficult to describe the behavior of players in more detail because of the many selection processes they could possibly use.

formula that accounts for the bias introduced by conscious selection, and, as opposed to Farrell and Walker, get more significantly different demand parameters when compared to the case of random selection.

Regarding the empirical analysis of players' behavior we try not to waste information from the already limited data set by looking at lower winning classes also. Compared to Scoggins (1995) and Papachristou and Karamanis (1998) we can thus come up with more reliable evidence on the question how lotto players do predictions. The result of those papers, namely that random selection is to be rejected, is however supported.

The structure of the paper is as follows: in Section 2 we formally introduce the notion of a parimutuel lottery and discuss the behavior of rational participants in more detail than in the previous literature, investigating to what extent the assumption of random selection is justifiable. After introducing the data set in Section 3 we derive the price of a ticket as a function of observable variables like the number of participants and the size of the rollover based on independent random prediction selection. Using the Austrian data from 1986–97 (weekly drawings), we estimate a demand curve in Section 4. We test for deviations from the rational players' selecting behavior in Section 5. Finding that these deviations are large enough to affect the quantitative results, we re-estimate demand properties under an alternative, empirical, pricing function in Section 6.

2. MODELLING PARIMUTUEL LOTTERIES

2.1. Background

A general parimutuel lottery is a mechanism that reallocates funds amongst the people participating in it according to a set of predefined rules. The operator has the duty to execute the reallocation and gets in turn a fixed percentage of the total handle. All participants are free to choose any non-negative multiple of a minimal monetary contribution to the total. Call the minimal contribution the bet b, measured in monetary units. The individuals contributing a positive amount are called the participants or players, indexed $i = 1, \ldots, I$. Any contribution of bx purchases x tickets. The total number of tickets is denoted T. Each ticket confers the right (and in fact the duty) to make a prediction about the outcome of the *draw*. The draw is a realization of a uniformly distributed random variable. The fact that all draw outcomes are equally probable distinguishes parimutuel lotteries from other parimutuel gambles as the football pools. This fact is taken as common knowledge to all players, unlike the predictions of other players, which are private information. The draw is usually organized the following way: from a total of N balls, labeled $1, 2, \ldots, N$, n balls are drawn without replacement. We say that a ticket is a *winner of class m* if its prediction, which in our case involves again n numbers taken from $1, 2, \ldots, N$, matches m of the numbers actually drawn.[5] The total of bets, the pool, is then redistributed in the following way: the operator receives a fixed share upfront. This share is called the take-out $h \in [0, 1)$. The rest is divided amongst classes $m = 0, 1, \ldots, n$ according to the shares s_m with $\sum_{m=0}^{n} s_m = 1$. Finally, any class pool $s_m bT(1-h)$ is distributed amongst the players proportionally to their holdings of winning tickets of class m. If no winning ticket of class m exists, then the class pool of class m is not distributed amongst players but retained to be distributed in the next round amongst winners of the same class m. The amount retained from previous rounds is called *rollover* of class m, short R_m. The values of b, N, n, h and s_m are stated in the rules and constitute the game design. They are important for both, the attractiveness of the lottery as well as for the expected loss from purchasing a ticket.

For future reference, we table the game design of the Austrian national lottery. A ticket wins in class 5b if it matches 5 of 6 numbers plus an additional, separately drawn, number. A ticket wins in class 5 if it matches 5 of 6 numbers and it does not match that additional number. b has been 8 Schilling (approximately €0.58) since 1991, before it was 6 Schilling (€0.43).

Table 1 The Austrian Lotto 6/45

Variable	Value
b	8 ATS
N	45
n	6
h	0.563125
s_0	0
s_1	0
s_2	0
s_3	0.25
s_4	0.20
s_5	0.15
s_{5b}	0.10
s_6	0.30

2.2. Rational players' behavior

Assume that the utility an individual derives from gambling in the Lottery $L(b, N, n, h, s; R)$ depends on the number of tickets purchased only. Given that an optimal amount is wagered, we can ask the question: what predictions are maximizing his payoff? Assume that the individual is risk-neutral.[6]

5. The game is often more complicated with separate classes for bonus balls etc.
6. As we do not motivate the decision to gamble by risk attitude but by the presence of an intrinsic utility of gambling, this is not a critical assumption.

Contrary to what might be common belief, the decision on the numbers to be predicted is not irrelevant even in a gamble with pure random draw. The structure of the lottery renders this decision to be a strategic one, as, unlike fixed-odds games, the decision of other participants will influence the own payoff. An instructive example is an event that occurred in the German Lotto 6/49. In April 1999 the numbers 2, 3, 4, 5, 6 and 26 were drawn. This resulted in a stunning 38,008 winners in class 5, each receiving a meager 380DM. The usual size of the prize in this class is 20 to 40 times higher. To see the point most clearly, consider the following example.

Example: A fair coin is tossed. There are two risk-neutral players that participate by wagering \$1 each and predict the outcome of the toss (H for Heads or T for Tails). The predictions are done simultaneously. There is no take-out, so the pool of \$2 will be shared by the winners. Should there be no winner, the pool is lost for the players, as in the case of a rollover.[7]

This is a parimutuel lottery in the above sense with $b = 1$, $N = 2$, $n = 1$, $h = 0$ and $s_1 = 1$. The normal form of this game is given in Figure 1.

This game has three Nash equilibria: (H, T), (T, H) and a mixed strategies equilibrium with $\sigma^1(H) = 0.5$, $\sigma^1(T) = 0.5$, $\sigma^2(H) = 0.5$, $\sigma^2(T) = 0.5$ where $\sigma^i(S)$ denotes the probability of player i choosing the pure strategy S. In the mixed strategies equilibrium, players would thus flip a coin in the first place to determine their own prediction.

The mixed strategies equilibrium has convenient symmetry properties. We state now that to mimic the actual draw is a Nash equilibrium in the general case if we retain one assumption. Regarding the probability to win in the given set-up, notice the following well-known statement:

Lemma. The probability of a given prediction to be a winner of class m, p_m, in a lottery $L(b, N, n, h, s)$ is given by

1 \ 2	H	T
H	-0.5, -0.5	0, 0
T	0, 0	-0.5, -0.5

Figure 1 Normal form of a simple parimutuel lottery

7. Of course, in the case of a rollover, the players will benefit in the next round, but we assume throughout that they will not take this into account. In the real-world lottery this makes sense because the own stake will have a negligible effect on the expected ticket loss in future rounds as the number of participants is huge.

$$p_m = \frac{\binom{n}{m}\binom{N-n}{n-m}}{\binom{N}{n}} \tag{1}$$

Proof. Straightforward; as drawn balls are not replaced, the number of matches follows the hypergeometric distribution.

Let

$$\mathcal{N} = \binom{N}{n}$$

denote the number of possibilities to draw n balls out of N without replacement.

Given that players $j \neq i$ choose the strategies $[\sigma^j(k)]_{k=1,\ldots,\mathcal{N}} = p_n$, the number of (other) winning tickets is a binomially distributed random variable, such that the payoff of i choosing *any* pure strategy is given by

$$\Pi^i = -b + \sum_{m=0}^{n}[bI(1-h)s_m + R_m]p_m \sum_{r=0}^{I-1}\frac{1}{r+1}\binom{I-1}{r}(p_m)^r(1-p_m)^{I-1-r} \tag{2}$$

where I is the number of participants. The fact that this payoff is independent of the chosen prediction proves:

Proposition. In a parimutuel lottery $L(b, N, n, h, s; R)$ with I participants submitting one ticket each, the strategy combination $[\sigma^i(k)]_{k=1,\ldots,\mathcal{N}} = p_n$ $\forall i = 1, \ldots, I$ constitutes a Nash equilibrium.

Although this seems quite natural, note that the equal objective probabilities in the drawing are essential for this symmetry argument. To see that the argument does not carry over to a more general case of 'favorites' and 'outsiders' as occurring in parimutuel sports betting, modify the above example as follows: suppose that instead of the coin toss, a ball is drawn from a box containing two red and one black ball. The task is to predict the color of the ball drawn. Contrary to what might be believed, the strategy combination where all players mix according to the probabilities $P(\text{Red}) = 2/3$, $P(\text{Black}) = 1/3$, will not be a Nash equilibrium. Given that one player does this, the expected payoff for the other of choosing Red is $(2/3)[(2/3) + 2(1/3)] - 1 = -1/9$ which is strictly larger than $(1/3)[(1/3) + 2(2/3)] - 1 = -4/9$, the expected payoff of choosing Black.

The above equilibrium can serve as a game-theoretic foundation of the common ad hoc assumption in the previous literature that all players select their predictions independently and with equal probability, which is the prerequisite for deriving the formula for the implicit price of lottery participation. Our framework challenges this assumption. First, the symmetric equilibrium is clearly not unique, which was already illustrated by the example before. In fact, in the I players, $\mathcal{N}$ draw outcomes game, there is a huge number of equilibria involving pure strategies choices. They even are Pareto-superior to the symmetric one. To see this consider a lottery with $I = \mathcal{N}$ where clearly a situation with all number combinations being selected by unique players is a Nash equilibrium in pure strategies (as in the coin-tossing example). There will never be a rollover here and thus all players are better off in terms of expected payoff. Also other equilibria with part of the players choosing a pure strategy and others mixing are possible. This relates to a criticism frequently found in the literature (see for instance Scoggins, 1995). It is argued that 'conscious selection', such as playing numbers forming a pattern on the betting slip or covering just the numbers from 1 to 31 (which appear in birth dates) is commonplace, causing an underestimation of the implicit price of lottery participation, as in this case a rollover is much more likely to occur. Regarding people who stick, for whatever reason, to 'their' combinations as playing pure strategies and the rest choosing at random, but avoiding the popular combinations, the probability of a rollover could actually fall, leading to a lower implicit price than predicted. This reasoning is not that far off reality as it may seem. In countries with well-established lotteries, usually professionals can be found, who, often in the form of a syndicate, would investigate the behavior of players in recent drawings by comparing the number of winners with the numbers drawn and then bet on other combinations. This behavior incorporates a positive externality by reducing the probability of a rollover. Because of these phenomena, it is impossible to tell with certainty in which direction the estimated price will be biased and ignorant participants will not drive the price up provided that enough rational and informed players are present. In his empirical work, Simon (1998) finds that one needs to distinguish between at least three groups of players in order to explain the combinations actually chosen well: those that play popular numbers, those that play random numbers and also those that, probably with a strategic motivation, go for unpopular numbers.

Assuming that all players are rational and selecting popular 'lucky numbers' does not improve on the utility derived from taking part in the gamble, the selection of the symmetric equilibrium seems most reasonable. The game we look at here has $\mathcal{N} = 8{,}145{,}060$ and the number of tickets usually surpasses 13 million, submitted by an estimated 1.5 million players.[8] These huge numbers make coordination between players almost impossible.

Besides the equilibrium selection, there is a minor problem with the assumption on players' behavior. Participation can involve submitting several tickets. It is impossible to give an exact average number because participation is anonymous, but market research of the Austrian Lotteries suggests that the average player submits around eight predictions per drawing. In this case, selecting predictions independently from each other will be suboptimal, as, by accident, two or more predictions may be 'near' to each other and may win in the same class. Consider the two valid predictions 1, 2, 3, 4, 5, 6 and 1, 2, 3, 4,

8. Notice that Austria's population is only 8 million.

5, 7 that could be the result of randomly and independently selecting two predictions. Should the numbers 1, 2, 3, 4, 5, 8 be drawn both win in the class of five matching numbers. Any equilibrium involving some form of coordination is Pareto-superior to the symmetrical one. While we reject pure strategy equilibria because coordination between different players may not be feasible, it is certainly possible for *one* player to coordinate her predictions. Optimally, the predictions should not have more than two numbers in common, three matches being the lowest winning class. To get an idea of the magnitude of the loss, when selecting predictions independently, consider the example of a player submitting two tickets.[9] When choosing independently, the probability of m matches on the tickets is the same as that a given ticket wins in class m, as in the above lemma. Given a number of other class m winners W_m the difference in expected payoffs of the optimal versus the random strategy is then

$$2bT(1-h)\sum_m p_m^2 s_m\left(\frac{1}{W_m+2}-\frac{1}{W_m+1}\right) \tag{3}$$

Plugging in parameter values from our actual sample, this amounts to $b(-5)10^{-8}$. The difference in payoffs is therefore just $(-2.5)10^{-6}$% of the stake $2b$. Note that the adoption of the truly optimal strategy requires quite an amount of sophistication and effort in filling in the ticket forms. Meanwhile the proposed equilibrium strategy of imitating the actual drawing sufficiently often in an independent way is extremely easy to handle for the players. In 1994, during the period covered by our sample, the Austrian Lotteries introduced the so-called 'Quicktip' feature. It permits the participants to have their tickets filled in automatically, with the combinations being drawn at random from a uniform distribution. Not only is this service widely used today (1994: 11.5%; 1996: 19.3% of all tickets), it is precisely the behavior predicted in the symmetrical equilibrium.[10] As the difference in expected payoffs is small for the lottery we look at, introducing boundedly rational players in the sense that pursuing any sophisticated playing strategy carries a small cost would eliminate this difficulty.[11]

9. With just two tickets the payoff difference is easy to calculate. More tickets involve very tedious combinatorics. The related problem of how many tickets one must submit in order to guarantee a win in any class is still unsolved to our knowledge.
10. Before the introduction of Quicktips, players may have made a pre-drawing on their own using widely available computer programs or drawing machines similar to the one used by the Austrian Lotteries themselves. They could even buy tickets already manually filled in for no extra charge.
11. Oddly, it is common in Austria that players accept even *worse* than selecting their predictions independently. There are betting slips that permit to submit multi-predictions in abbreviated form, for instance to predict 7 instead of just 6 numbers, which translates into making 6 single predictions. It should be clear from the discussion that this behavior is inferior in expected payoff terms, but again, with a large pool, the difference will be small and could be rationalized by the ease of filling in and checking these forms.

3. THE DATA SET: AUSTRIAN LOTTO 6/45

Lotto '6/45' was introduced in Austria in 1986. It has immediately become the most popular gamble in Austria as measured by the share of the population taking part. Its present share in the Austrian lottery market (which includes also scratchcard games, a fixed-odds numbers game and two parimutuel football betting games) is 51%. Although after the introduction of 6/45 severe substitution effects could be observed – the turnover of the football pools decreased by nearly 50% – more than about 80% of the turnover of 6/45 in 1986/87 was new money not been spent on gambling before. Lotto 6/45 is run by the Austrian Lotteries, a private company which holds the only license to offer large-scale lottery games in Austria. During our sample period from September 1986 to August 1997 there have been weekly drawings each Sunday, giving 570 observations in total. Our sample ends in August 1997, as in September 1997 a second drawing on Wednesday was introduced leading to a structural break in the data. The Wednesday drawing caused a decrease in participation on Sundays of approximately 25%. However, Wednesday drawings are still less popular than Sunday drawings. The average number of bets on Wednesdays is about 28% lower than on Sundays. Still, total sales increased after the introduction of the second drawing. Given a population of 6.65 million over 15, the average Austrian has wagered 19 Schilling (€1.38) per round in 1996, which amounts to 991 ATS per year. In 1998 average annual expenditures on 6/45 rose to 1,281 ATS (€93) and expenditures per round decreased to 12.4 ATS. In 1998 total expenditure on 6/45 amounted to 8.53 billion ATS which equals 0.325% of the Austrian GDP. Lotto 6/45 is also a non-negligible tax source. Thirty-eight per cent of the wager goes directly to the state. The tax receipts from 6/45 account for 0.8% of total federal tax revenues. The value of the aggregate pool has been at least 50 million ATS for each Sunday drawing since 1994. In drawings with three subsequent rollovers the size of the jackpot has amounted to about 200 million ATS. About 22 million lottery tickets win prizes each year.

The full data set we used is downloadable from the internet.[12] Alongside the numbers drawn, it includes the number of winners and the size of the prize in each class. Knowing the values of h, s and b given in the previous section, the value of T can be readily calculated from this. There are just a few problems which were accounted for. First, the data give no hint on possible additional augmentations to promote sales. This would result in a miscalculation of T, but we retrieved the missing information.[13] Second, the prizes are rounded off to the next full schilling. This makes the take-out rate slightly higher, as the operators profit from each round will increase by around 250,000 ATS. Third, during our sample period, in 1991, the stake b was increased from 6 ATS to

12. http://www1.lottery.co.at/docs/index2/html
13. Occasionally, non-monetary promotion prizes were offered, like a car. These were not accounted for.

8 ATS. As to be explained, this led us to split the sample in two subperiods with 253 observations for the 6 ATS regime and 316 for the 8 ATS regime.

4. EXPECTED LOSS AND DEMAND

We use the theory on players' behavior from Section 2 to devise a formula for the expected loss per ticket, q. We saw in (2) that under the condition that each player chooses predictions independently from a uniform distribution, this expected loss amounts to

$$q = b - \sum_{m=0}^{n} p_m[bT(1-h)s_m + R_m] \sum_{r=0}^{T-1} \frac{1}{r+1} \binom{T-1}{r} (p_m)^r (1-p_m)^{T-1-r} \tag{4}$$

Using the fact that T is large while the p_m are small, we can invoke the Poisson approximation to the binomial distribution to get

$$q = b - \sum_{m=0}^{n} p_m[bT(1-h)s_m + R_m] \sum_{r=0}^{T-1} \frac{1}{r+1} e^{-p_m(T-1)} \frac{[p_m(T-1)]^r}{r!} \tag{5}$$

or

$$q = b - \sum_{m=0}^{n} \frac{[bT(1-h)s_m + R_m]e^{-p_m(T-1)}}{(T-1)} \left(\sum_{r=0}^{T} \frac{1}{r!} [p_m(T-1)]^r - 1 \right) \tag{6}$$

As before, the large number of tickets allows the approximation

$$\sum_{r=0}^{T} \frac{\lambda^r}{r!} \cong e^{\lambda} \tag{7}$$

Thus we get

$$q = b - \sum_{m=0}^{n} \frac{(1 - e^{-p_m(T-1)})(bT(1-h)s_m + R_m)}{T-1} \tag{8}$$

or, if the effect of the own ticket is ignored,

$$q = bh - \sum_{m=0}^{n} \left(\frac{R_m}{T} - e^{-p_m T} \left(b(1-h)s_m + \frac{R_m}{T} \right) \right) \tag{9}$$

Notice that the expected loss per ticket in a given Lotto round is not just the take-out, bh. Rollovers from previous rounds increase the expected payoff.

Theoretically, it is possible for the ticket value to exceed the stake, provided that a sufficiently high rollover combines with a low participation. We shall see that, in the sample we use, this never happened. There are additional loss terms, however, that arise because of the possibility that a rollover happens in the current round. The size of this loss is determined by the value of $p_m T$. The higher this expression is (easy game design combined with a high participation) the less likely is a rollover, resulting in a lower implicit ticket price.

Other articles (for instance Cook and Clotfelter, 1993) assume that for $m < n$, the expression $e^{-p_m T}$ is small enough to be ignored. Because a rollover in class 5b (five out of six plus additional number) has happened twice in our sample we will include the rollover factor for that class in our pricing equation to take the following final form:

$$q = 0.563125b - \left(\frac{R_6}{T} - e^{-(T/8145060)}\left(0.1310625b + \frac{R_6}{T}\right)\right) - \left(\frac{R_{5b}}{T} - e^{-(T/1357510)}\left(0.0436875b + \frac{R_{5b}}{T}\right)\right) \tag{10}$$

where we plugged in the values for $p_6,\ p_{5b},\ s_6,\ s_{5b}$ and h for the Austrian Lotto 6/45 from Table 1.

Walker (1998) already noted that the sign of dq/dT is ambiguous in the presence of a rollover, while it is clearly negative when no rollover has happened in the previous round. We have determined the magnitude of the elasticity of q with respect to T for all sample lotto rounds. While in rounds with no previous rollover $\varepsilon_{q,T} \equiv (dq/dT)(T/q)$ is almost constant at -0.07, we find that it is positive for *all* rounds with a previous rollover, with an average of 0.117 and a standard deviation of 0.091 (146 observations). The lowest value was 0.042 and the highest a striking 0.498, implying that, in that round, a 1% increase in sales would bring about a 0.5% increase in the expected loss of a ticket. This confirms Walker who asserted that, because sales in rollover rounds are already high, the loss term accruing from the smaller rollover share per ticket would dominate the benefit of a larger pool, which reduces the probability of another rollover happening. This contrasts sharply with the advertising policy of the Austrian Lotteries, which emphasizes the sheer size of the jackpot especially in rollover rounds, because, as demonstrated, a higher total jackpot would actually increase the implicit price of participation.

We are now interested in a demand specification for Lotto tickets. Figure 2 shows the number of tickets sold in each lotto round for the whole sample period. The huge increases in demand in weeks with a double or triple rollover are clearly visible.

Figure 3 shows a plot of T versus the implicit price of a ticket q for the full sample period (570 Lotto rounds from 1986 to 1997). First, observe the accumulation of observations around 10 million tickets at a price of 3.5 ATS

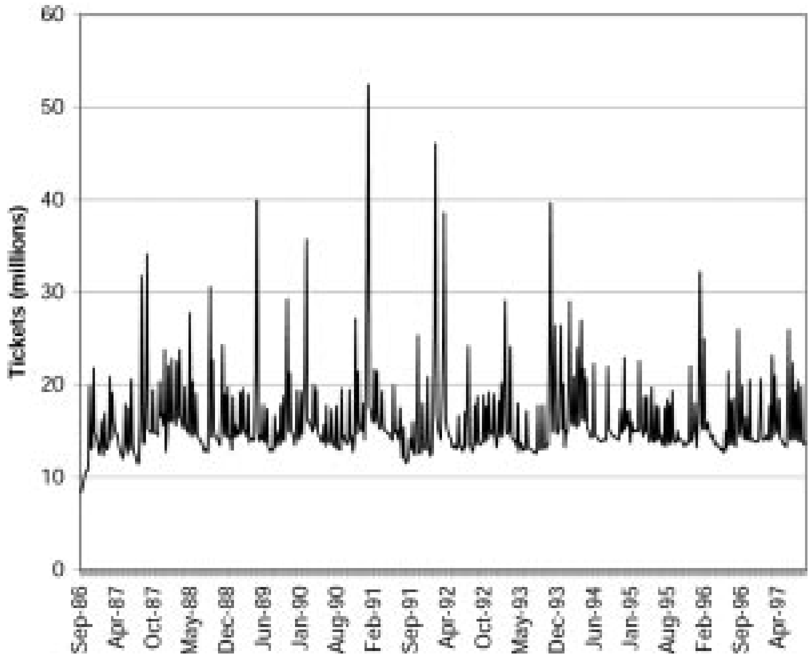

Figure 2 Ticket sales September 1986 to August 1997

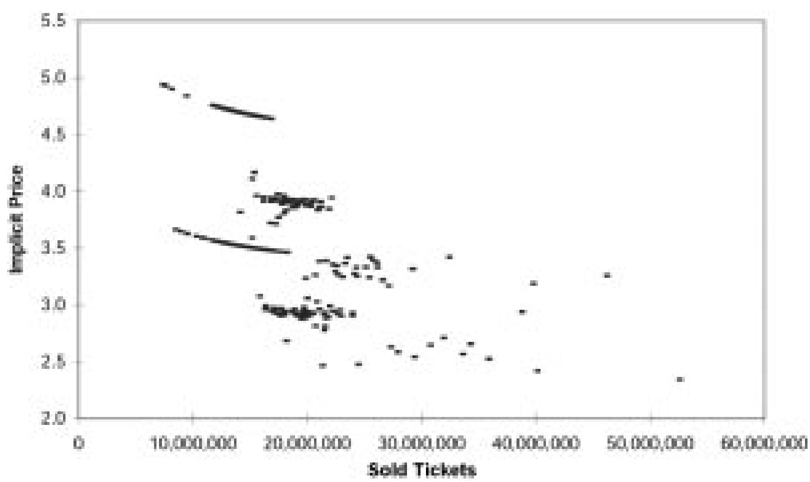

Figure 3 Participation and the implicit price

and at 4.5 ATS. These observations correspond to rounds with no rollover in the period with $b = 6$ and $b = 8$, respectively. In these cases the relation between price and quantity is governed entirely by the pricing schedule (10), which the model assumes to be deterministic. As the only predetermined variables are the rollovers R_6 and R_{5b}, these observations do not contribute to

Table 2 Average price and demand: rollover rounds and non-rollover rounds

	Mean sales	Mean implicit price
Regular rounds	14 million	3.53 (6 ATS period)/4.80 (8 ATS period)
Rollover rounds	21.3 million	2.86 (6 ATS period)/3.73 (8 ATS period)

determining demand parameters. The scattered observations elsewhere correspond to rounds where a previous rollover introduces exogenous variation.

Second, a structural break in the data occurring in 1991, when b rose from 6 ATS to 8 ATS is highly visible. The rise in b, where at the same time h remained constant, induced a nearly proportional increase in the expected loss of a ticket, which our model assumes to be the main variable influencing demand. The empirical fact that demand was nearly unaffected by that rise cannot be reconciled with such a simple view. Apparently such a major change in game design as a variation in b changes consumers' propensity to pay (in expected loss terms), as the two gambles are viewed as different. It is however impossible to verify empirically a more complicated model accounting for this, as the changes in b are much too rare. A similar problem occurs in Papachristou and Karamanis's (1998) work on the Greek Lotto 6/49. Their resort of using the expected ticket value per monetary unit, q/b, is theoretically unfounded. We will therefore split our sample into 253 observations with $b = 6$ and 316 with $b = 8$ such that all observations in a subsample have the same game design.

We propose a linear demand schedule:

$$T(q) = d + kq \tag{11}$$

The functional form of the demand specification will not be too important, as the data do not allow for more than an estimation of local properties. While a constant-elasticity demand function may appear more natural at first sight, it is likely that a linear specification is a better choice for the unobservable range of prices. It can be expected that, even at an expected loss of zero, the number of tickets submitted is finite, because there is a limited number of individuals that derive utility from gambling at all and the marginal utility of submitting one more ticket will eventually become zero when the total stake is high enough as compared to the prizes offered. On the other hand, there is a natural upper bound on the implicit price, which is the stake b. A lottery with an expected loss higher than or equal to b cannot offer a positive payoff at any positive probability, which constitutes a fundamental property of gambles.

The following system of equations is thus estimated:

$$T_t = d + kq_t + u_t \tag{12}$$

$$q_t = 0.563125b - \left(\frac{R_{6t}}{T_t} - e^{-(T_t/8145060)}0.1310625b + \frac{R_{6t}}{T_t}\right)$$
$$- \left(\frac{R_{5bt}}{T_t} - e^{-(T_t/1357510)}\left(0.0436875b + \frac{R_{5bt}}{T_t}\right)\right) \quad (13)$$

Tables 3 and 4 give the full information maximum likelihood estimations and 95% confidence intervals for d and k measured in millions for the two subsamples. We excluded one observation from the second subsample, namely round 37 in 1996. In this round there has been an exceptional promotion to celebrate ten years of 6/45. All prizes were calculated as usual but then doubled up. The take-out rate in this round was just $1 - 2(1 - 0.563125) = 0.12625$, compared to 0.563125 in all other rounds, resulting in an implicit price of just 1.2 Schilling per ticket. Oddly, only around 20 million tickets were sold, which is strikingly low compared to round 4 in 1996, for instance, with more than 32 million sold. In spite of the double rollover in that round, ticket purchase was still associated with an expected loss of 3.4 Schilling. This observation would underline the importance of the distribution of the pool amongst classes for the intrinsic utility of a gamble, because rollovers will bias this distribution towards the grand prize class while in the promotion round the distribution was the same as in non-rollover rounds.

Note that the estimate for k is significantly lower in the $b = 8$ regime then with $b = 6$. Translated into elasticities, we get -1.7 for rollover rounds under $b = 6$, where the implicit price was around 2.9. In the $b = 8$ period an elasticity of only -1.3 is obtained (implicit price in rollover rounds is around 3.7). Whether this is due to the unmodeled aspects of the changing game design or some exogenous change in behavior of participants is hard to tell.

Ordinary least squares estimations for the parameters do not differ significantly. The R^2 for the OLS estimation is 0.698 for the period with $b = 6$ and 0.671 for the period with $b = 8$. We also wanted to check the influence of additional possibly relevant variables. Other authors have proposed alternative

Table 3 Estimation results for $b = 6$

Parameter	Estimate	Lower limit	Upper limit
d	61.1	57.6	64.6
k	−13.4	−14.5	−12.4

Table 4 Estimation results for $b = 8$

Parameter	Estimate	Lower limit	Upper limit
d	49.8	47.6	52.0
k	−7.7	−8.2	−7.2

models. Farrell *et al.* (1999) investigate the impact of lagged previous rollovers on the demand for lottery tickets and find that sales rise not only in a rollover week, but also in the week after. So we also tested for the presence of serial correlation and estimated

$$T_t = d + kq_t + mL_t + u_t \tag{14}$$

where the additional term L_t is the number of weeks that passed since the last time a rollover occurred (not taking into account the running week's rollover, if there is one). If there is some myopic addiction factor determining lotto sales, the coefficient m should take a negative value. We find a negative and significant coefficient for both periods. However, the values of the estimates for m are relatively small, pointing to a negligible impact of the term L_t and, more important, the R^2 increases from 0.671 to 0.692 and from 0.698 to 0.739 only, revealing the tiny additional explanatory power of lagged rollovers.

5. ARE PREDICTIONS SELECTED RANDOMLY?

The analysis of the previous section rests on the condition that lottery participants choose number combinations randomly and independently. While we could show that this behavior could be interpreted as a Nash equilibrium, there is considerable empirical evidence against it.

First, the results from the analysis of numbers actually selected in the German and the Swiss Lotto are in contradiction with the random selection assumption. Bosch (1999) reports for the German Lotto 6/49 significantly favoured and neglected numbers. He finds further that numbers are not equally distributed across columns and rows and neighbouring numbers are less frequently chosen than expected under the hypothesis of random selection. A non-negligible share of lotto players also prefers 'nice looking' patterns on the lottery ticket, birthday dates or winning combinations from previous draws. An analysis of the Swiss Lotto by Riedwyl (1990) or Henze and Riedwyl (1998) and the results of Simon (1998) for the UK lotto yield similar results. Ziemba *et al.* (1986) also found popular and unpopular numbers and combinations for the Canadian Lotto. Although the data for the Austrian Lotto does not allow for an investigation on the level of individual tickets,[14] we must expect a similar pattern governing the behavior of Austrian Lotto players.

The alternative to analyzing the – for the researcher usually unavailable – data on the individual predictions is the statistical analysis of the outcome of lottery draws: winning numbers, the number of winners per class and the distribution of jackpots. An outcome as in round 20 in 1989, where 23

14. In Hauser *et al.* (2000), we examine the Austrian football pools ('Toto'), having detailed micro-data for 166 rounds available and find the same myopic patterns in a gamble that shares many properties with Lotto.

participants won the first prize in the Austrian Lotto, makes the random selection assumption appear unlikely. Under this assumption, the probability of this event is about $6(10^{-18})$, plugging in the value of T for that round. Other authors report similar evidence for other countries' lotteries. Scoggins (1995) investigated the Florida Lotto. Under the null hypothesis of random selection, the probability of a grand prize rollover is

$$P = (1 - p_6)^T \tag{15}$$

He extends (15) arbitrarily to

$$P_t = (1 - p_6)^{\alpha + \beta T_t} \tag{16}$$

where $P_t = 1$, if a rollover occurred in round t, and $P_t = 0$ otherwise. Using maximum likelihood, he estimates (standard errors in parentheses) $\alpha = 8.75$ (3.84), $\beta = 0.261(0.229)$ and the null that $\alpha = 0$ and $\beta = 1$ can be rejected. Papachristou and Karamanis (1998) repeated this for the Greek Lotto 6/49. They arrive at $\alpha = 6.7(3.22)$, $\beta = 0.431(0.105)$ and reject random selection as well.[15]

Applying Scoggins's method to the Austrian data, we get $\alpha = 2.25(3.14)$, $\beta = 0.573(0.207)$. Although β is higher than in the other studies and not significantly different from 1, the joint hypothesis that $\alpha = 0$, $\beta = 1$ is still rejected by a standard likelihood ratio test.

It is doubtful that the above procedure is the most useful to check whether players select numbers randomly (but we propose a different use for it later). A better method is to test the implication of random selection instead, which is that the number of winners is Poisson (or binomially) distributed with parameter $\lambda_m \equiv p_m T$ where m is a class with $s_m > 0$. As p_m is fixed and T known for each Lotto round, the theoretical distribution of winners is known for each round and each class. Scoggins's method uses just a small part of this information, namely the probability of a rollover in the highest class. While it is straightforward to test for any single round, the aggregation is problematic. Simply adding up the number of winners and participants over the rounds and using the reproductive property of the Poisson distribution is not appropriate as unusually high and low numbers of winners would even out in total.[16]

More reliable evidence is provided by the following method: for large sample sizes,

$$Z = \frac{(W_m - p_m T)}{\sqrt{p_m T}} \tag{17}$$

15. They estimate (16) using least squares, which is methodologically questionable.
16. Indeed, the data do not invalidate the hypothesis that the aggregate number of winners in any class over the 570 rounds in the sample is Poisson distributed with parameter $p_m \sum_t T_t$.

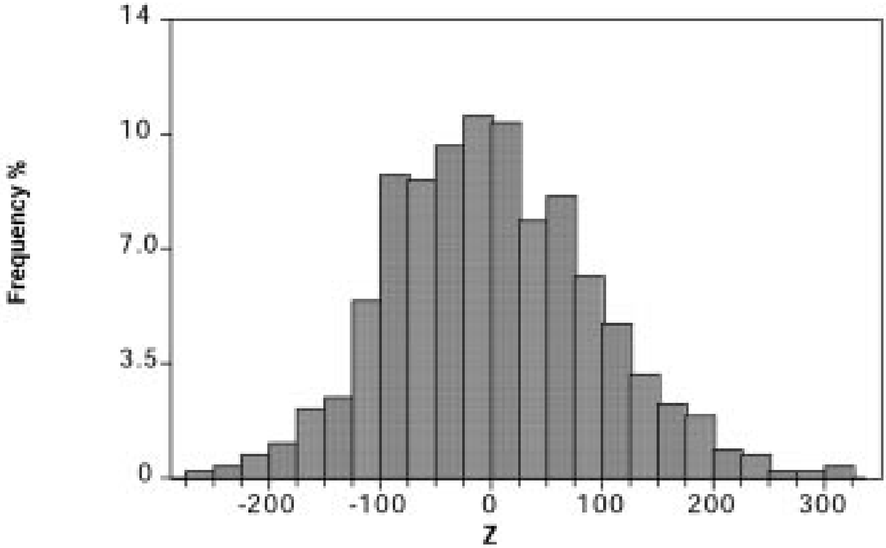

Figure 4 Frequencies of empirical Z-values

should follow a *standard* normal distribution. Taking the empirical values for $W_{3t},\ t = 1, \ldots, 570$ for instance, we get the following frequencies for Z_t plotted in Figure 4.

From Figure 4 it is evident that Z does not follow a standard normal distribution. The variance is about 100, with a maximum of 314.6 in round 51/1987, and a minimum of −232.1 in 51/1989. Class 3 winners in the latter round received almost three times as much as in the former. A decreasing variation in these values would point towards some learning amongst participants in the sense that predictions become more evenly distributed over time. The development of the variation over time can be analyzed by dividing the total sample of Z values into ten overlapping subsamples of equal size. The window width of each subsample is 103 with an overlap of 50 realizations. Figure 5 shows that the variance of the subsamples follows a decreasing trend from 1988 onwards. The decrease in variance becomes most obvious in the subsamples containing 1994 and later years. We explain this by the introduction of the 'Quicktip' in 1994.

A different indicator for non-random selection by players was suggested by Haigh (1997). If all players predicted the draw outcome independently, then no positive correlation between the proportions of winners in separate class pools would be expected. Indeed, as every ticket can win just in one class, one would expect a high proportion of players winning in class m to be accompanied by a slightly lower proportion winning in class $n \neq m$ in that round; so, if anything, the correlation should be negative. However, if some numbers, or combination of numbers, are more popular than others, a positive correlation can result. Table 5 gives correlations for the 570 rounds from the Austrian 6/45.

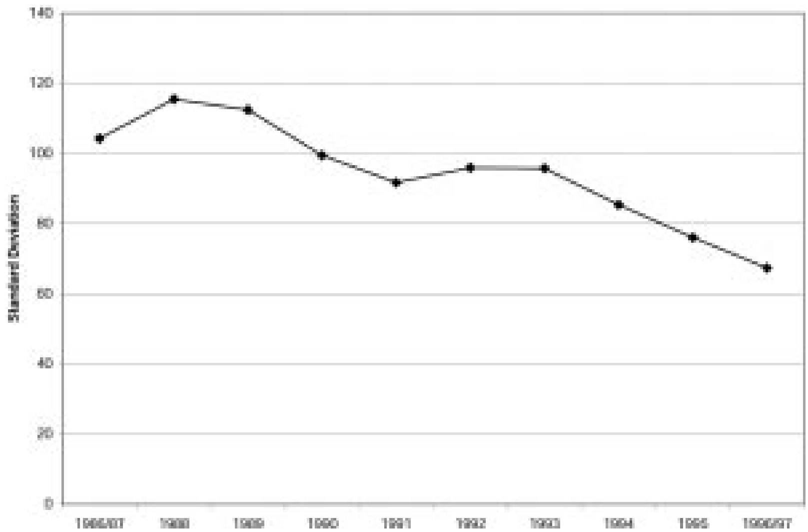

Figure 5 Standard deviation of Z over time

Table 5 Correlations of proportions of winners in different classes

Class	5b	5	4	3
6	0.40	0.55	0.48	0.41
5b	–	0.49	0.52	0.47
5		–	0.79	0.67
4			–	0.92

These high positive correlations do not only indicate that some combinations are vastly more popular than others with an insufficient number of players pursuing a strategy of betting on unpopular numbers. They can also be a result of the widespread use of abbreviated multi-prediction forms, which typically result in either no win or multiple wins in different classes, because the use of these forms leads to strongly correlated individual predictions. In that case, the users of these forms, who usually bet more than the average player, seem not to be much better in selecting optimal predictions than others.

6. AN ALTERNATIVE MODEL

Given the evidence of conscious selection, i.e. players choosing predictions non-randomly, the calculation of the implicit price as done in Section 4 may not be correct any longer. While the qualitative properties of the pricing function will continue to hold when allowing for more general forms of players' behavior (for instance, the price in a non-rollover round will usually still go down with T), the magnitudes may well differ. When the advantage of the symmetry is lost, however, analysis becomes substantially more difficult. First, different predictions will have different prices and would have to be treated as heterogeneous goods if a full-information assumption is to be maintained. This is clearly intractable, also in view of the available data. Second, even when continuing to assume that all players use the same prediction strategy, it is unclear what such a strategy should look like. Farrell and Walker (2000) assumed that players assign different probabilities to the N numbers, not necessarily equal to the true probabilities in the actual drawing. These probabilities can be inferred from the data about drawn numbers and sizes of prizes. Leaving aside the statistical uncertainty involved, this ignores the important role of popular combinations, like patterns on the betting slip.

We propose a simple, but still reasonable, approach instead. The expected price of participation basically depends on three factors: (a) the take-out rate, (b) rollovers from previous rounds and (c) the possibility that part of the prize pool is not distributed due to a rollover. The first two factors are exogenous for each round, while the latter is sensitive to the way people choose predictions. Assume now that players engage in conscious selection, i.e. they make suboptimal predictions because they lack either information or sophistication, or both. Then, they are also not in a position to calculate an expected loss of their bet, not knowing the exact probability of a rollover happening. The only way for them to assign a price to a lotto ticket is therefore to use information about past drawings to infer this probability. The simplest way to do this would be to calculate the fraction of rounds in the past where a rollover happened and take that value (around 25.6% in our sample) as fixed. A slightly more sophisticated player would, however, learn that the rollover probability is related to the number of participants T in each round. The precise relationship is unknown because the participants are unable to process information about the precise nature of conscious selection. Thus it seems natural to employ the simple extension

$$P = (1 - p_6)^{\alpha + \beta T} \tag{18}$$

with α and β estimated from the data as done in the previous section. As we saw, the case $\alpha = 0$, $\beta = 1$ corresponds to independent random selection. While here the parameter test is probably a bad indicator for whether people select randomly because a lot of information is unused, in the present context it is appropriate because rollovers of the grand prize are what determines the implicit price. Also rollovers in class 5b can be considered; however, in our sample they happened just twice which is insufficiently often to determine coefficients for α and β here. The effect of a changing probability is negligible here anyway. We will instead use a fixed probability of 2/570 for rollovers in that class.

Table 6 Estimation results for $b = 6$

Parameter	Estimate	Lower limit	Upper limit
d	73.4	70.3	76.6
k	−15.9	−16.8	−15.0

Table 7 Estimation results for $b = 8$

Parameter	Estimate	Lower limit	Upper limit
d	60.4	58.2	62.5
k	−9.3	−9.7	−8.8

The modified price equation for the Austrian Lotto 6/45 thus becomes

$$q_t = 0.563125b - \left(\frac{R_{6t}}{T_t} - e^{-(\alpha+\beta T_t/8145060)}\left(0.1310625b + \frac{R_{6t}}{T_t}\right)\right) - \left(\frac{R_{5bt}}{T_t} - \frac{2}{570}\left(0.0436875b + \frac{R_{5bt}}{T_t}\right)\right) \quad (19)$$

where $\alpha = 2.25$ million and $\beta = 0.573$ are the maximum likelihood estimators from Section 5. We can repeat estimating (12) and (19) simultaneously by FIML to get the results (in millions) shown in Tables 6 and 7.

For the relevant range of T, the modified rollover probability is always higher than in the random selection model. Also, an increase in T will reduce the probability of a rollover by less. Therefore, under the conscious selection model, the expected loss of participation will be always higher than under random selection and the difference widens as T goes up. To explain the given data set, the demand parameter k must be higher in absolute value then. When we translate the results into elasticities again (evaluating at a typical rollover round), the results of about -1.9 (for the $b = 6$ period) and -1.5 (for $b = 8$) signify a higher elasticity under both regimes compared to the results under the standard random selection model.

7. CONCLUSIONS

Games of chance where money is at stake have persisted throughout history. Economists have been puzzled by the wide acceptance of these – often unfair – gambles by individuals that seem to be risk-averting elsewhere. The easiest way to reconcile the empirical facts with the theory of consumer choice is to assign to the process of gambling an intrinsic utility. Keeping in mind how dangerous such an approach is – any economic model can be fitted to reality in this fashion – it is evident that the circumstances under which money is put at risk play a role for many consumers. To engage in risks in the casino or at the races seems to be much more pleasant than to recall the possibility of your house burning down. This may still be true for some individuals even if the same amounts are at stake.

It is then natural to ask how much people are willing to pay for a given type of gamble. Given that take-out rates differ a lot across gambles, the whole set of game characteristics will influence demand, yet, the expected loss of taking part should play an important role. By looking at only one type of gamble, the other characteristics are held fixed. The problem of insufficient variation of the take-out rate in a given gamble is overcome by considering rollovers in parimutuel lotteries. While this is an interesting and fairly unique approach, it is also a problematic one. First, intertemporal substitution may be important. If Lotto players tend to shift stakes from non-rollover rounds to rollover rounds, the estimated demand parameters must be interpreted with care. A decrease in the take-out rate, which affects the long-term expected loss, may not lead to the full boost in demand as predicted, because elasticities tend to be biased downwards (in absolute value) if there is intertemporal substitution.

Second, rollovers do not only affect the expected value of a ticket, but also higher moments. Because rollovers typically concern the grand prize, the distribution of the pool amongst classes changes. This means that the game design is not really fixed in this respect. Unfortunately, it is impossible to account for these simultaneous changes because of an identification problem. We saw that a promotion of the Austrian Lotteries, where not only the grand prize but all classes were affected, becomes an outlier when considering just the price as an explanatory variable, which underlines the importance of the distribution amongst classes. Assuming that increasing only the grand prize pushes ticket demand more raises the question whether the current class shares s_m are optimal, however.

Last, the implicit price of playing the lottery heavily depends on the way people do their predictions. The usual assumption that they choose at random is sustained as a Nash equilibrium. There are other equilibria involving coordination amongst players, which result in a lower expected loss per average ticket, but our evidence supports the results of other authors, namely that people do even worse. Assuming random selection, one would expect around 87 rollovers to have happened in the history of the Austrian Lotto, while actually 146 happened. This not only drives up the average price of playing, it also means that an additional participant who selects consciously, has less of a positive externality on others than assigned by the random selection model. While our results indicate that there is some learning in making predictions, even after more than ten years, the number of prize winners varies much more than expected under random selection.

Because of conscious selection, it may be more appropriate to use the empirical model we present in Section 6 for the purpose of determining the optimal take-out rate of a given lottery.

ACKNOWLEDGMENTS

The authors wish to thank Michael Boss, Elisaveta Krylova, Raimund Alt (all Institute for Advanced Studies, Vienna) and Egbert Dierker (University of Vienna) for valuable comments and suggestions.

REFERENCES

Bosch, K. (1999), *Lotto und andere Zufälle*, Oldenbourg.

Boss, M., B. Felderer, U. Hauser, Ch. Helmenstein and U. König (1997), *Das österreichische Glücksspielmonopol*, Institute for Advanced Studies, Vienna.

Clotfelter, Ch. and Ph. Cook (1989), *Selling Hope: State Lotteries in America*, National Bureau of Economic Research, New York.

Conlisk, J. (1993), 'The Utility of Gambling', *Journal of Risk and Uncertainty* 6(3), 255–275.

Cook, Ph. and Ch. Clotfelter (1993), 'The Peculiar Scale Economies of Lotto', *American Economic Review* 83(3), 634–643.

Farrell, L. and I. Walker (2000), 'The Welfare Effects of Lotto: Evidence from the UK', *Journal of Public Economics* 72, 99–120.

Farrell, L., E. Morgenroth and I. Walker (1999), 'A Time Series Analysis of UK Lottery Sales: Long and Short Run Price Elasticities', *Oxford Bulletin of Economics and Statistics* 61, 513–526.

Farrell, L., R. Hartley, G. Lanot and I. Walker (2000), 'The Demand for Lotto: The Role of Conscious Selection', *Journal of Business and Economic Statistics* 18(2), 228–241.

Gabriel, P. and J. Marsden (1990), 'An Examination of Market Efficiency in British Racetrack Betting', *Journal of Political Economy* 98(4), 874–885.

Haigh, J. (1997), 'The Statistics of the National Lottery', *Journal of the Royal Statistical Society* 160, Part 2, 187–206.

Hauser, U., U. König and E. Krylova (2000), 'Gamblers' Rationality in Parimutuel Soccer Betting', Institute for Advanced Studies Working Paper No. 87.

Henze, N. and H. Riedwyl (1998), *How to Win More – Strategies for Increasing a Lottery Win*, Wellesley, MA.

Papachristou, D. and D. Karamanis (1998), 'Investigating Efficiency in Betting Markets: Evidence from the Greek 6/49 Lotto', *Journal of Banking and Finance* 22, 1597–1615.

Riedwyl, H. (1990), *Zahlenlotto – wie man mehr gewinnt*, Bern.

Scoggins, J. (1995), 'The Lotto and Expected Net Revenue', *National Tax Journal* 48(1), 61–70.

Simon, J. (1998), 'An Analysis of the Distribution of Combinations Chosen by UK National Lottery Players', *Journal of Risk and Uncertainty* 17(3), 243–276.

Suits, D. (1979), 'The Elasticity of Demand for Gambling', *Quarterly Journal of Economics* 93(1), 155–162.

Thaler, R. and W. Ziemba (1988), 'Parimutuel Betting Markets: Racetracks and Lotteries', *Journal of Economic Perspectives* 2(2), 161–174.

Walker, I. (1998), 'The Economic Analysis of Lotteries', *Economic Policy* 27, 357–401.

Ziemba, W., S. Brumelle, A. Gautier and S. Schwartz (1986), *Dr. Z's 6/49 Lotto Guidebook*, Dr Z Investments, Vancouver.

