The Economic Journal, **111** (*November*), F700–F722. © Royal Economic Society 2001. Published by Blackwell Publishers, 108 Cowley Road, Oxford OX4 1JF, UK and 350 Main Street, Malden, MA 02148, USA.

AN ECONOMIST'S GUIDE TO LOTTERY DESIGN*

Ian Walker and Juliet Young

This paper outlines the issues relevant to the design of pari-mutuel lottery games and makes inferences about game design effects from estimates of how rollovers affect sales. Lottery tickets sales depend positively on the proportion of revenue returned as prizes, positively on the *skewness* of the prize distribution (which depends largely on how much of the prize money goes to the jackpot), and negatively on the *variance* in the prize distribution (which depends largely on how much goes on smaller prizes). We simulate the effects of envisaged game design changes on sales revenue and find potentially large effects.

The intention of this paper is to step back from the controversy and to consider how the UK on-line pari-mutuel lottery game (called 'Lotto' in the United States and elsewhere) *ought* to be designed, operated, taxed and regulated.[1] Little attention has so far been given to the considerations raised here and yet they are central to both the objectives that government have set for the operator and for wider objectives such as the welfare of society as a whole. Unfortunately, there has also been little analytical research into how lottery games work, or how they should be operated and designed.

There are statistical issues concerned with how to structure the game to generate sales and this involves choosing the number of combinations of numbers that can be bought in a way that makes the game attractive to players. The statistical design affects how hard it is to win, and this affects how attractive the game is on a draw-by-draw basis, and so also affects how sales might be expected to behave in the long run. There are economic issues concerned with the sizes of the prize pools for different winners. The bigger the overall prize pool, the better the bet being offered and the more attractive the game will be. Market structure is also important: on-line pari-mutuel games exhibit economies of scale – bigger games are more 'efficient' than small ones (in a very specific sense to be defined later) so regulating entry into the market is likely to be very important. Moreover, the stability of sales is likely to be adversely affected by a competitive market structure – one important feature of pari-mutuel games is that rollovers, which are random events, cause the attractiveness of tickets to change and if several suppliers are offering near-identical products the demand for each will be unstable. Of course, monopolistic supply will generally imply a need for effective regulation – and this will be true even if the licensee is operating on a not-for-profit basis. Finally, game design can be fine tuned to exploit the preferences of players so as to improve sales. For

* This paper draws on Ian Walker's research with colleagues at Keele University that was funded by the ESRC under research grant R000236821, and on Juliet Young's MSc dissertation at the University of Warwick. The views expressed here are those of the authors alone and do not represent the views of our employers.

[1] See Munting (1994) for the history of UK lotteries.

example, the distribution of the overall prize pool between jackpot winners and lesser prize winners may affect the attractiveness of the game.[2]

1. Statistical Considerations in Lottery Game Design

On-line[3] games usually feature players buying tickets where they choose n numbers from a possible N available numbers. Such games are usually *pari-mutuel* in design – that is winners, whose tickets match the winning combination (or some part of it), receive a *share* of the prize pool with any other players who match the same number of numbers. The chances of winning depend on n and N – the bigger is n, the harder it is to match the winning numbers since you have to match more of them, and the bigger is N the more possible combinations there are to be matched. Part of the problem of designing such a game is one of choosing the right vales of n and N for the market circumstances. If $n = 6$ and $N = 49$ then the probability of a ticket being the winning combination is (approximately) 1 in 14 million, while if $n = 6$ and $N = 53$ then the chance of buying the winning combination is (approximately) 1 in 23 million. Thus n and N affect the likely number of winners: with $n = 6$ and $N = 49$ and 60 million tickets sold then the number of jackpot winners to be expected is more than 4.3, but if $N = 53$ then the expected number of winners is less than 3. These are the mean numbers of winners that we would expect – there is a variance around these numbers and the implications of $N = 53$ rather than 49 is an increase in the chance of there being *no* winners. In the event of there being no winners, the jackpot is added to the jackpot of the next draw – this event is known as a rollover.

The behaviour of sales over time rests largely on the choice of n and N, which determine the probabilities of winning the different prizes and the likelihood of a rollover. If the game is easy to win, then rollovers are infrequent so each draw is much the same as the next and there is a danger that players become bored with the monotony of the game. Estimates in an earlier paper see Farrell *et al.* (1999) suggested that the 'half-life' of the UK game would have been approximately 150 draws – sales would halve every 3 years (of weekly draws) – if there had been no rollovers. Rollovers enhance the attractiveness of the *next* draw so that players are enticed to then play more, come back to

[2] There are three areas where we have little to say. First, technology affects both how games can be presented to players and the kind of game that it is possible to organise. Pari-mutuel games that allow players to choose their numbers have been spawned by the availability of sophisticated computer systems. But new technology also offers the prospect of internet-based games and games operated via mobile phones using SMS or WAP. The technological possibility of international competition also imposes constraints on the domestic market as well as offering further market possibilities. Second, gambling can have adverse social consequences and intelligent game design can be used to minimise these. However, imposing constraints on game design because of a concern over adverse social consequences will generally have adverse consequences for sales so a trade-off may be involved. For example, it might be regarded as better to have a large number of small players than a small number of large ones. Finally, scratch cards are a part of the portfolio of the UK game and we have little to say about this since we do not have good data for them.

[3] In the lottery industry, 'on-line' means games where ticket sales are recorded electronically at a dedicated terminal.

the game, or join the game for the first time, and this effect takes some time to decay.

However, a game that is too hard to win will also be bad for sales in the long run. In the extreme case, imagine a game that was almost impossible to win, it would rollover almost forever since sales would be very low and hence few of the available combinations would be bought in any draw. But the size of the rollover would very slowly accumulate and hence so would sales. This is an example of 'intertemporal substitution' – players sit on their hands waiting for the jackpot to grow sufficiently large for the draw to become attractive and, only then, play heavily. Even in less extreme cases, rollovers give rise to intertemporal substitution since rollover draws are more attractive than regular draws. While it is true that extra sales occur when there is a rollover, this is, in part, at the expense of sales in regular draws. Thus, designing the game to maximise sales is a balancing act of making it hard enough to win to overcome the tedium but easy enough to win to avoid significant intertemporal substitution.[4] Thus, it is important that the game design matches the likely size of the market: a game that is sensible for the UK is likely to be too hard for Israel whose population is just 10% of that of the UK. In fact, the Israeli on-line lotto game was recently redesigned from 6/49 (1 in 14m) to 6/45 (1 in 8.1m) precisely because the operators felt that it was too difficult to win and rollovers were too frequent. In contrast, the game in Ireland (population 3.8m) has twice been redesigned to make it harder to win to induce more rollovers. Indeed, the redesigns followed organised attempts to 'buy the pot' because large jackpots had accrued. Under the new design, a 6/42 game so that the odds of winning are 1 in 5.25m, there are more frequent but smaller jackpots. In California (population 34m), the game began as 6/49, went to 5/53 and then to 6/51 (1 in 18 million) but, since June, has a complex $5/47 + 1/27$ design that gives extremely long jackpot odds of 1 in 41.4m. In Florida (population 15m), the game has also recently become more difficult, going from 6/49 to 6/53.

The prize pool is defined by the take-out rate, τ, which is the proportion of sales (ie the stakes) that is not returned as prizes. Thus, the overall prize pool is $(1 - \tau)S$, where S is sales revenue (in many games the cost of a ticket is fixed at a unit of currency so S is both the number of tickets sold and the level of sales revenue). It is common for the take-out rate to be in the range 40–50% so that the pay-out rate is 60–50%.[5]

Smaller prizes are usually awarded for matching fewer than n numbers, so it is common for the prize fund to be split into separate pools. More complex designs are possible – for example, in the UK Camelot game, there is a seventh

[4] The problem is made more complex where there are other substitution possibilities – for example, in the United States, it is possible that cross-state substitution takes place. This gives rise to incentives for neighbouring states to collude and share the proceeds of a single large game rather than have two competing games.

[5] Care must be taken when comparing across games to recognise that some games pay prizes as a lump sum (in the UK, for example) while others (eg most US states) pay an annuity (or some heavily discounted lump sum). Moreover, in some countries (eg the United States) the prizes are liable for income tax while in other countries (e.g. UK) they are not.

'bonus' ball that is also used to define a prize pool for matching 5 of the first 6 numbers drawn plus the bonus number. Thus, the overall prize pool is usually divided into separate pools for funding players who match all n numbers in the winning combination, match $n-1$, $n-2$, etc. This set of prize pools might be characterised by $s = s_1, s_2, \ldots, s_n$. In the Camelot game $s_1 = s_2 = 0$ and s_3 is not a share at all but a fixed payout and these match-3 prizes are awarded first and the shares of the other prizes is defined out of the residual.[6]

The odds of matching fewer than all n numbers also depends on N: thus the odds of matching 3 in a 6/49 design is 1 in 57, while the odds of matching 3 in 6/53 is 1 in 71.[7] Thus both n and N affect the number of prize winners for each prize pool, and it is the shares (and τ) that affect the amount of money in each prize pool. Thus, the average amount won by each type of prize winner depends on all of the design parameters of the game. So, for any specific n and N, the design of the distribution of the prize money, through choice of τ and the s_i shares, affects the mean return from buying a ticket (which is less than $1-\tau$ because of the rollover probability – the higher the rollover probability, the lower the return *to the current draw*), the variance in returns around this mean, and the degree to which the prizes are skewed towards large or small prizes. The larger the share given to the jackpot, the more positively skewed is the distribution of prizes and the larger the share given to the lower prizes, the more negatively skewed is the distribution.[8] The variance depends on how much weight is given to middle as opposed to extreme prizes. Note that these 'moments' of the prize distribution are not independent of each other: for example, reallocating the prize money away from the easy-to-win prizes and towards the hard-to-win prizes increases the skewness but also increases the variance. It will also lower the mean return because a higher jackpot share will imply that a larger proportion of the revenue staked will be *at risk* of rolling over (for a given rollover probability).[9]

One way of summarising the complications of how all the various aspects of game design impacts on sales is through the mean, variance and skewness of the prize distribution. However, estimating the empirical impact of these three moments of the prize distribution on sales is difficult. One might be tempted to conduct an experiment where the design features were changed and sales

[6] That is $s_i = p_i[(1-\tau)S - 10N_3]$, where $i = 4, 5, 5+b, 6$, N_3 is the number of players that match 3 of the numbers drawn, and p_i is a fraction. For example $p_6 = 0.52$.

[7] Gerry Quinn in Ireland provides a helpful website that allows probabilistically challenged readers to compute the odds for many common game designs. See http://indigo.ie/~gerryq/Lottodds/lotodds.htm.

[8] Games that are hard to win often feature large jackpot shares. For example, in the Florida on-line twice weekly lotto draw the odds of matching 5 of the 6/53 has a (relatively) high chance but it has such a small share of the overall prize pool that it is only, on average, worth approximately $5,000. That is, the Florida lotto game is both hard to win and highly skewed. It is the large jackpot that entices people to play in regular draws even though there is a high chance that it will be rolled over and won by someone in subsequent weeks.

[9] The industry view is that the low value and high odds prize pool serves the role of ensuring that players and potential players are frequently reminded of the possibility of winning so that most potential players will know a recent winner. In fact, we find no evidence that the lagged number of match-3 (£10) prize winners, which has surprisingly high variance, has any statistically significant effect on sales.

recorded or, even, to offer one group of individuals one game design and a control group another design. However, *pari-mutuel* games exhibit increasing returns so the behaviour of a small treatment group would be of little value for estimating how such a game would sell if offered to the population. In practice, such experiments are not available to us and we observe either no variation in game design over the history of sales or, at best, changes in game design that the operator has chosen with a view to increasing sales. That is, any variation in game design that we may observe in *any* dataset is unlikely to tell us anything useful about, for example, how sales would change if a policy maker wanted to change the tax rate levied on the game.

In practice, the best we can do is to try to make inferences about how sales would be affected by game design changes, from the *random* variation in the terms on which people participate that we can typically observe – that is, through the effect of variations in the size of the jackpot on sales. The size of the jackpot is a random variable because rollovers are statistically random events. The value of each of the moments of the prize distribution depends on the game design parameters and on the level of sales – for example, for any given design, the mean return on a ticket is higher, the higher are sales. Thus rollovers cause there to be exogenous variation in the nature of the prize distribution.

Fig. 1 shows the 'expected value' of a lottery ticket for common types of design in a regular (non-rollover) draw. Expected value is the average return to buying a £1 ticket.[10] In the figure, the take-out rate τ is set at 0.55 which is a

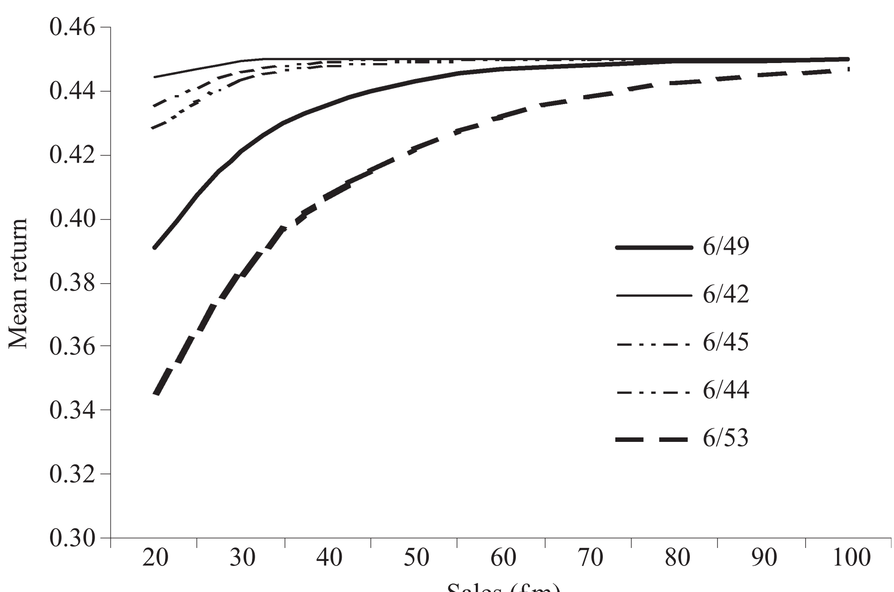

Fig. 1. *Lotto's Peculiar Economies of Scale*

[10] Tickets are commonly available in one unit of the local currency.

typical value and approximately the value used in the UK lotto game. The shape of this figure has given rise to what has been called lotto's 'peculiar economies of scale' since it shows that the game becomes cheaper to play (in the sense that the expected loss is smaller) the higher are sales. See Cook and Clotfelter (1993). This is because: the higher are sales the smaller is the chance of a rollover occurring because more of the possible combinations are sold;[11] this makes the return higher in the current draw because rollovers take money from the current draw and add it to the next draw; and your ticket in this draw gives you a possible claim on prizes in this draw but not the next. So, the higher the chances that a jackpot rolls over, the less a ticket for the current draw is worth. Note that, at very large levels of sales, all games have the same mean return which simply equals $1 - \tau$, because the chance of a rollover is small when ticket sales are large since most possible combinations will be sold. Notice also that, at any given level of sales, easier games offer better value in regular draws since the rollover chance is smaller.

Fig. 1 shows the situation for regular draws. However, when a rollover occurs, the mean, variance and skewness all change, and the way in which they change can be calculated from a knowledge of the determinants of these moments. In Fig. 2 we show, for a 6/49 design, how mean, variance and skewness vary with sales and how these relationships are shifted when there is a small rollover (£4m) and a large rollover (£8m). Rollovers make a difference in kind to the relationship between the moments and the level of sales for the following reason. In regular draws, players simply play against each other for a slice of the overall prize pool which comes from stakes in the current draw – since players play against each other, any addition to the prize pool is matched by additional potential winners. But, in rollover draws, players are also playing for the jackpot pool from the previous draw[12] and the value of this extra is spread more thinly as more tickets are sold. Thus, the relationship between sales and the mean of the prize distribution is made up of what would happen in a regular draw (as shown in Fig. 1) that would have the upward sloping economies of scale characteristic *plus* the value of the previous jackpot which falls as sales rise because its value is spread more thinly the more players are competing for this fixed sum. Thus, overall, as the top panel of Fig. 2 shows, the expected value first rises (as the economies of scale effect dominates) and then falls (as the competition for the fixed rolled-over amount takes over and the economies of scale effect flattens out).

The probability distribution implied by the UK 6/49 prize structure has a large spike at −£1, since mostly players lose, and a further smaller spike at £10, where 1 in 57 tickets match 3. For the pari-mutuel prizes, however, it is more difficult to describe the rest of the distribution, which is associated with the more difficult-to-win prizes, because the amount won depends on the number

[11] The rollover probability is $(1 - \pi_6)^S$ where π_6 is the jackpot odds (1/14m in the 6/49 case) and S is the level of sales.

[12] In principle, lower prize pools could also roll over but we have no evidence that this has ever occurred in practice.

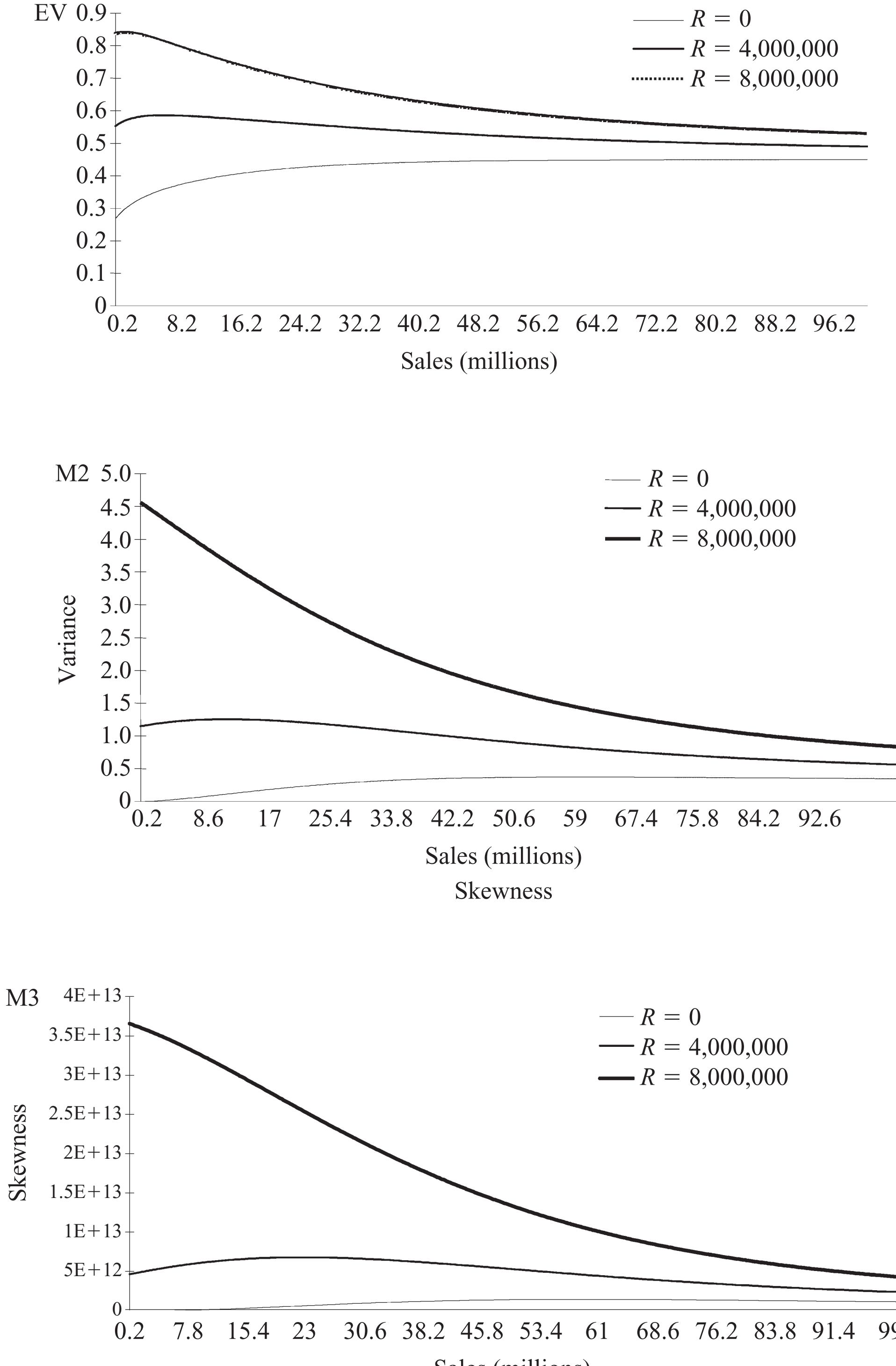

Fig. 2. *The Relationship Between Moments and Sales*

of people who also win a share in each prize pool. Instead of a spike, there is a small peak with a (local) maximum in the distribution corresponding to each prize type with a spread around the probable number of winners for that type caused by the draw-by-draw variance in the number of winners. That is, when unpopular numbers are drawn, there will be fewer winners each with a larger share of the pool, and when popular numbers are drawn, there are more winners than average each with a smaller share. Successive peaks, corresponding to the mean winnings of bigger prizes are lower (as the chance of winning is smaller) and wider (because the variance in the number of prize winners is higher for the more difficult to win prizes). The overall distribution is thus left skewed. The bottom two panels in Fig. 2 show how the variance and skewness of the prize distribution vary with sales and the rollover size. A rollover decreases the left skew (ie it increases (right) skewness) since it increases the size of the jackpot pool.

Fig. 3 shows the effects of rollover size on the mean, variance and skewness for two levels of sales, typical of Wednesday and Saturday draw levels. A rollover affects only the top prize and increases (right) skewness.[13] Increasing ticket sales has no impact on the two mass points corresponding to winning nothing or £10 but increases the prize pool for the other prizes and also the likely number of winners. With no rollover, the first effect dominates see Clotfelter and Cook (1991) and the increase in sales increases the expected value and the peaks of the distribution corresponding to the higher value prizes move rightward. With a rollover, however, the second effect dominates for high sales, and, although the expected amount won for the 4, 5 and 5+bonus prizes increases, the expected amount won in the jackpot prize may decrease.

Table 1 shows the actual values of the moments for typical examples. The message is that rollovers have a large effect on the mean, variance and skewness of the prize distribution, especially at low levels of sales, while the effects of variation in sales (for a given rollover size) is relatively small, especially at large levels of sales.

2. Previous Empirical Research

Few previously published papers have looked at the modelling of lotto sales. Scoggins (1995), a US example, suggests that *decreasing* the takeout rate for the jackpot prize could *increase* revenues (for the Florida state lottery).[14] The decrease in takeout rate would have two opposing effects: it would decrease revenue since, *ceteris paribus*, less money is taken as profits, but the larger prizes made possible would increase sales and thus increase the 'tax' revenue raised. While the increase in sales would also decrease the probability of a rollover, the increase in the probable size of any rollover which does occur more than makes up for this.

[13] We ignore the probability of lower prizes experiencing a rollover since this is very small.
[14] See Forrest *et al.* (2001) for UK work that follows this line but does not support the proposition.

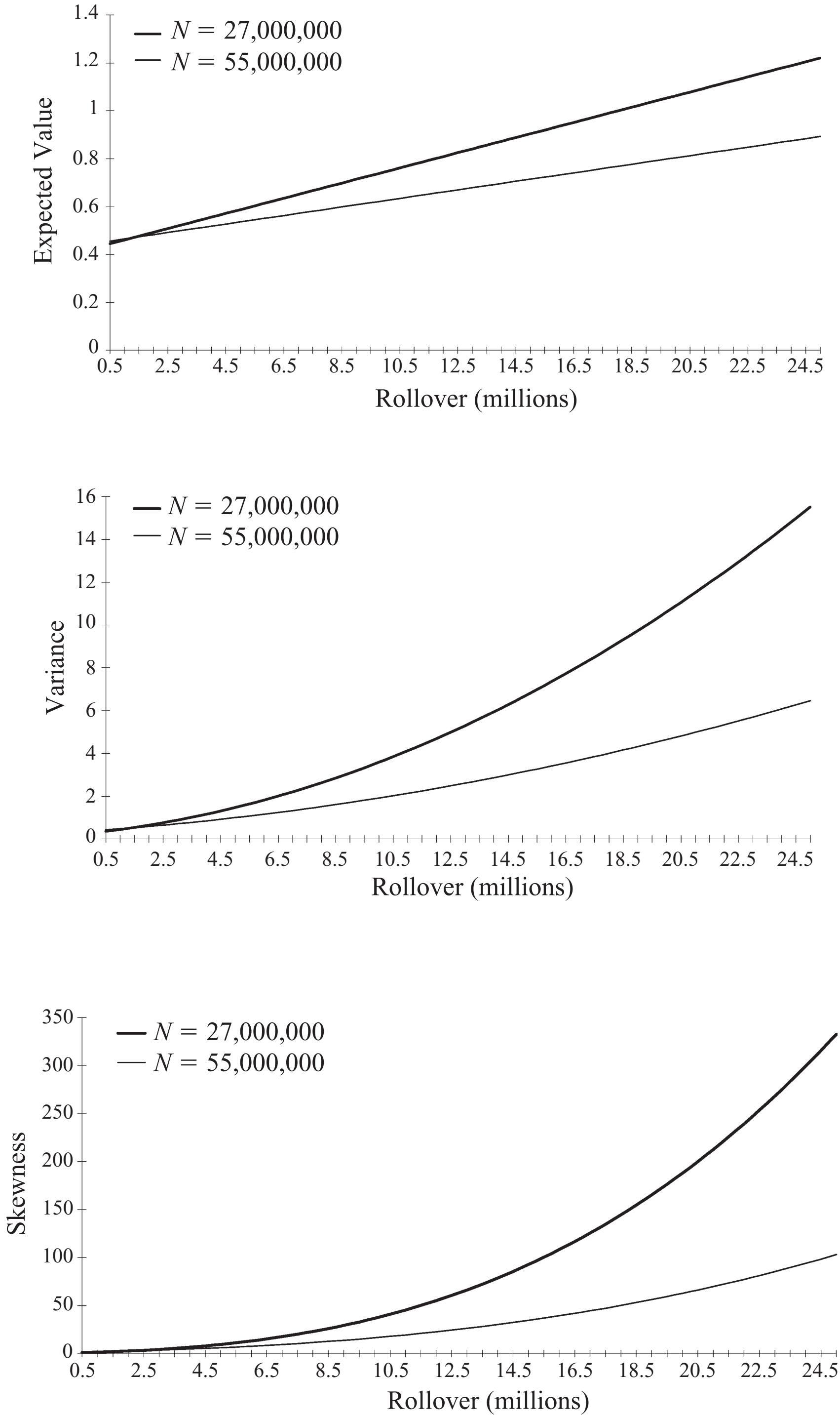

Fig. 3. *The Relationship Between Moments and Rollover Size*

Table 1

Variation of Moments with Changes in Sales and Rollover Values

(a) Typical Saturday sales				
Sales (millions)	60	80	60	80
Rollover (millions)	0	0	10	10
Mean	0.4480	0.4495	0.6124	0.5741
Variance	373817	362262	1738480	1262490
Skewness	1.4126×10^{12}	1.2896×10^{12}	14.3768×10^{12}	8.5044×10^{12}
(b) Typical Wednesday sales				
Sales (millions)	30	40	30	40
Rollover (millions)	0	0	10	10
Mean	0.4333	0.4418	0.7276	0.6775
Variance	303475	350841	3324420	2623430
Skewness	0.9136×10^{12}	1.2257×10^{12}	33.8585×10^{12}	25.4434×10^{12}

To consider the overall effect of a change in the takeout rate, Scoggins estimates two equations: one showing how sales depend on the size of the jackpot; and the other showing how the rollover probability is affected by any change in sales. Then, for a probability of winning the jackpot of π_6 the probability of no-one winning the jackpot must be $(1-\pi_6)^S$ if sales are at level S. There is, however, a tendency for players to choose numbers in a non-random fashion (a phenomenon termed 'conscious selection') so that some combinations of numbers are chosen more often than others. This implies that the 'coverage' of numbers is lower than would otherwise be the case and thus the rollover probability would be underestimated by the expression above. Instead, Scoggins assumes that the rollover probability can be represented by the form $(1-\pi_6)^{aS+b}$ and estimates the coefficients a and b to uncover the rollover probability. Using the rollover probability equation and the ticket sales equation, average sales and average tax revenue can be calculated as a function of the takeout rate, and the optimal takeout rate can be determined.[15]

Beenstock *et al.* (1999), using Israeli data, also considers the effects of changing the takeout rate. However, this work stresses the importance of rollovers in creating additional excitement and publicity which they refer to as 'lottomania'. The paper suggests that 'the optimal strategy consists of a delicate balancing act between increasing the incidence of rollover, since big money is made when lottomania takes possession of the public, and making it sufficiently attractive to play in the early rounds. If it were too difficult to win in the first round, there would be less money to be rollover over, and less lottomania. Again, they use the size of the jackpot, rather than the expected value, as the main determinant of sales.

The majority of the literature has thus been based on either the jackpot size

[15] By using the size of the jackpot rather than the expected value as the determinant of lottery sales, Scoggins overlooks the fact that a rational player would realise that on a rollover week, higher sales imply a smaller likely share in the jackpot if the winning number is chosen. The relationship between the jackpot and expected value therefore differs according to whether it is a rollover week.

or the expected value of the lottery.[16] Simple thought experiments, however, suggest that these do not capture the full effect that the distribution of prizes may have on demand. For example, if the *probability* of winning the jackpot and the existence of smaller prizes have no effect, then the use of the jackpot size as the sole explanatory variable begs the question of why lottery operators the world over bother with smaller prizes and why they fail to decrease further the odds of winning the jackpot in order to increase profits?

Moreover, empirical evidence of a preference for skewness in the distribution of prizes has been suggested by Golec and Tamarkin (1998) for racetracks, who suggest that the long shot bias, a commonly observed race-track phenomenon in which low-probability, high-variance bets (long shots) are overbet and favourites are underbet, can be consistent with risk averse behaviour if bettors have a preference for skewness. Garrett and Sobel (1999) take a similar approach to explain demand for (not necessarily pari-mutuel) lotteries. Both papers assume that preferences can be expressed as a function with prize money as the argument but the way in which the prize money enter the model is via three moments of the prize fund which, in their work, represent mean, variance and skewness, respectively. In both of these papers, the authors find that the coefficient on the first and third moment of the prize were significantly negative and the coefficient on the second moment was significantly negative, indicating a positive preference for money, an aversion to risk (variance), but a preference for skewness.

Garrett and Sobel (1999) assume that the lottery player's welfare depends only on the top prize payouts of each lottery game in each state. Where the prize structure was pari-mutuel, so that the top prize varied according to the number of tickets sold and the number of winners, the average top prize was estimated using annual sales and the takeout rate. Thus, the effects of game design on the skewness and variance of parimutuel lotteries are essentially ignored. In contrast to Garret and Sobel, in the work reported below, the mathematical mean, variance and skewness of the distributions are calculated and all used as explanatory variables. Although the dominant term in each of these variables are, respectively, the jackpot, the jackpot squared and the jackpot cubed (the variables used by Garret and Sobel), our approach also enables the effect of the other prizes and the level of sales to be captured.

Three recent UK papers – Farrell *et al.* (1999), Farrell *et al.* (2000), Farrell and Walker (1999) – present evidence on the determinants of sales in the form of statistical estimates of the extent to which sales increase for every £1 addition to the jackpot due to a rollover or superdraw.[17] These papers use a variety of datasets but are all couched in terms of rollovers affecting sales *only through the mean* return to buying a ticket. This work is extended and updated

[16] Forrest *et al.* (2001) attempts to test between these two competing hypotheses using UK sales data. They look at how sales vary with the jackpot size and compare this with expected value (ie the mean value of the prize distribution) which is a complex nonlinear function of the jackpot. However tests of nonlinearity will have little power when the data is strongly clustered around just two points (the mean of regular draws and the mean of rollover draws).

[17] We also deal with the distinction between superdraws that add to the jackpot and those that guarantee a minimum.

in recent work here to show how the effect of rollovers on sales can be decomposed into effects via the *mean, variance and skewness* in the prize distribution. Thus, we calculate the level of the mean, variance and skewness for each draw, allowing for the effect of the rollover size, and include these three variables into a time series model of sales.

3. Data and Estimates

The estimates, in Table 2, were obtained by simple least squares. Since our concern has been to identify how rollovers affect sales through the moments of the prize distribution and since rollovers are random events which are orthogonal to the other explanatory variables, we are not especially concerned with the time series properties of our model *per se*.[18] However, we are concerned with the endogeneity of the moments themselves because they all depend on sales (see Appendix for the definition of the moments). This is potentially problematic because there is no obvious way of instrumenting all of the moments – in Farrell *et al.* (2000), we include only the first moment and use the size of the rollover as an instrument. In fact, we found that OLS produced estimates that were not significantly different (as did nonlinear least squares which solved for the endogeneity) for the obvious reason that the mean return is very insensitive to the levels of sales at the typical level of sales. Thus, almost all of the variance in the mean return is due to rollovers. Inspection of Figs 2 and 3 show that this also holds for higher moments – at typical levels of sales, these graphs are close to horizontal. Thus, we rely here on least squares estimates.[19]

The results show that: sales are a statistically significant increasing function of the mean of the prize distribution – so better bets are more attractive ones; sales are a statistically significant but decreasing function of the variance in the prize distribution – so riskier bets are less attractive; and sales are an increasing

[18] Inspection of the residuals plots from the estimates in Table 2 suggests that the regression underpredicts sales for double rollovers and the first two double rollovers in particular. Although the AR test is passed, indicating absence of autocorrelation, as does the ARCH test, the χ^2 for heteroscedasticity and the normality and Reset tests fail. The heteroscedasticity test may be picking up genuine heteroscedasticity it can also fail because of some form of misspecification, which would mean that the coefficients are biased. However, plots of the squared residuals against the size of the jackpot enhancement seem to confirm that the regression is genuinely heteroscedastic and variance increases with the size of the rollover, which seems plausible. The normality test probably fails because of the large outliers caused by the first rollover which attracted considerable media attention. The failure of the Reset indicates the possibility of a misspecified functional form (although the log regressions does no better) or an omitted variable, either of which implies that the coefficients will be biased. Parameter constancy tests for both models were also carried out on the remaining observations and passed, although the forecast sample did not contain any double rollovers. The forecasts also systematically under predicted although this is probably more likely to be due to inadequate modelling of the time trend than the effect of rollovers since including the three moments already allows a good deal of flexibility even though there is relatively little variation in the sizes of rollovers.

[19] Although the numbers that players chose are non-random, the moments have been calculated on the basis of randomly chosen numbers, and so the expected value calculations will be biased downwards and, strictly speaking, differ with the choice of numbers. The results in Farrell *et al.* (2000) suggest that conscious selection has only a small effect on the mean.

Table 2

OLS Results: dependent variable lottery sales (millions)

	Coefficient	t-value
Constant	−91.92	−4.94
Mean	294.34	4.33
Variance ($/10^5$)	−4.17	−2.46
Skewness ($/10^{12}$)	1.63	1.83
Wednesday	−29.15	−4.70
Dummy for Sat. after intro. of Wed. draw	4.28	0.77
No. of terminals (1000s)	1.26	12.10
Weeks since first draw	−0.064	−10.26
No. of Wed. draws	0.165	6.96
Square of the Weds. Draw No.	−0.0005	−4.20
Thunderball dummy for Saturday draws	−0.923	−2.31
Thunderball dummy for Wednesday draws	0.098	0.16
Delayed Diana draw	−11.66	−15.23
Superdraw in which 5b prize topped up	−68.49	−4.65
1st lag of Sat. sales before intro. of Wed. draw	0.205	2.42
2nd lag of Sat. sales before intro. of Wed. draw	0.066	1.47
1st lag of Sat. sales after intro. of Wed. draw	0.149	2.01
2nd lag of Sat. sales after intro. of Wed. draw	0.023	0.63
1st lag of Wed. sales	0.028	0.53
2nd lag of Wed. sales	0.068	2.14
Quarter two dummy	2.434	4.20
Quarter three dummy	1.181	4.11
Quarter one dummy	1.276	5.02
R^2	0.973825	
Diagnostic Tests		
AR 1–2 F(2,447)	0.87574 [0.4173]	
ARCH 1 F(1,447)	0.20186 [0.6534]	
Normality $\chi^2(2)$	2607.2 [0.0000]	
χ^2 F(34,414)	17.403 [0.0000]	
RESET F(1,448)	111.36 [0.0000]	
Parameter Constancy		
Forecast $\chi^2(14)$	10.99 [0.6868]	
Chow F(14,449)	0.5078 [0.9289]	

Note: t statistics calculated using heteroskedastic consistent standard errors

(but bordering on significant) function of the skewness of the prize distribution – so players exhibit a (small) preference for skewness.

Sales exhibit statistically significant positive correlation across time – a rollover which raises sales in the current draw, will also raise sales somewhat in subsequent draws. This is termed the 'halo effect' in the industry and was captured in earlier work by Farrell *et al.* (1999). While it is true that we cannot distinguish between our interpretation of the results and a model which simply says that rollovers have a highly nonlinear effect of sales,[20] these same effects

[20] Camelot, on their website, state that 'after months of extensive research amongst the British Public it was found that the chance of winning millions was the most motivating strategy for potential British players'. No further details are given on their research methodology. Of course, if they really believed this then the lower prize pools would be better spent on the jackpot and they should redesign the pool sharing rule. However, in our analysis, the lower prizes have a role in promoting sales by reducing the variance that players dislike.

of mean, variance and skewness have been estimated before in the context of gambling on horse races in the United States and by comparing sales across US lottery games, and the same pattern emerged.

Several other explanatory variables were also used in the modelling. A variable was included to allow for differences in sales on Saturdays and Wednesdays.[21] A further variable was included to allow for a shift in Saturday sales once the Wednesday draw had been introduced to allow for possible substitution effects. The number of retailer terminals has also been included. This grew steadily for the first 110 draws and is intended to pick up a ease of availability. However, the change in the number of terminals is unlikely to account entirely for the rise in sales over the initial period and this variable may also be picking up the natural growth in sales following the introduction of a new product before equilibrium is established. This point is reinforced by the fact that the Wednesday draw also exhibits low sales initially. In the longer run, however, time tends to have a negative effect, as people become bored and lose interest in the game. To capture these effects, a trend representing the weeks since the introduction of the lottery was included. A separate linear and quadratic trend for the Wednesday draw were also included to capture change in interest for the Wednesday draw.[22]

Lagged variables to capture the effect of the previous draw were also included and are likely to reflect habit, or possibly addiction. They are also intended to pick up the halo effect which can be seen clearly in the graphs of sales; after a rollover has occurred, sales continue to be higher for the subsequent regular draws. The lag structure is likely to change following the introduction of the Wednesday draw so separate variables were used for lags before and after the introduction of the Wednesday draw. After the introduction of the Wednesday draw, lags may have a further importance, since it is possible to buy a ticket for the following Saturday draw when buying a ticket on a Wednesday and vice versa. Thus, higher sales on a Wednesday may imply higher sales on a Saturday because of the convenience of being able to buy tickets for Saturday's draw on Wednesday. Alternatively, there may be a negative effect if a Wednesday rollover induces people to substitute Saturday sales for Wednesday sales. Since the effect of Saturday draw on the following

[21] It could also be argued that the introduction of the Thunderball game may have changed the response to the moment distributions, since the consumers with less of a taste for skewness and a greater dislike of variance might have stopped buying the normal lottery tickets and, instead, bought tickets for the Thunderball draw. The hypothesis that the response to the moments was different after the Thunderball draw was introduced was tested by applying a Wald test to slope dummies corresponding to the introduction of the Thunderball draw. Although the hypothesis that these dummies were zero was rejected with the regression in its original form, producing an F statistic of 19.083 [0.0000], once Wednesday slope dummies had been introduced the hypothesis could not be rejected, producing an F statistic of 0.12067 [0.9479]. We also include a variable to capture the draw when the operator added £20m to the 5+ bonus prize pool.

[22] Variables were also included to correspond to the introduction of the Thunderball game in June 1999. This game has a far less skewed distribution, with a 1 in 4 million chance of winning the top prize of £250,000, and a 1 in 33 chance of winning the bottom prize of £5. Unlike the standard draw however, prizes are fixed in size and there are no rollovers. The game is drawn on a Saturday and generally averages sales of £4–5m. These sales may, in part, be taken from the main game, although it is notable that, in previous empirical work, little substitution between draws has been found, so it may be that the Thunderball draws in new sales, or alternatively, increases sales, by creating additional interest in the lottery.

Wednesday draw is not necessarily the same as the effect of a Wednesday draw on the following Saturday, lags of Saturday draws were entered separately from lags of Wednesday draws.[23]

It is possible that some of the specification problems arise due to the failure to include a variable relating to advertising and media coverage. Double rollovers attracted particularly large amounts of coverage during the early days of the lottery and this may explain the large outliers associated with the double rollovers. Modelling the inevitable rise and decline of interest in the lottery with a simple quadratic or linear time trend is almost certainly inadequate, although if this is uncorrelated with any of the other variables in the regression, this should not cause any bias.

Clearly, our estimates rely on the functional form restriction that rollovers affects sales only through their effect on the first three moments for the prize distribution. However, it is difficult to test this specification against alternatives because rollovers have rather limited variance so that tests of functional form would have little power. Thus, in the absence of other evidence, and with support from studies of gambling on other datasets, we feel that our interpretation of the estimates of the way in which sales vary with rollover size is the best available and we exploit the estimates to simulate the effect of game design induced changes in moments below.

4. Simulation

In principle, the estimates imply that we can make inferences about game design changes. Table 3 shows the theoretical distribution of prizes under the current arrangements and, for one particular draw, the actual distribution. The expectation in this particular draw was for there to be 4.5 jackpot winners sharing the jackpot pool of close to £8m. In fact, by chance, there were just two winners who each received close to £4m. Other prizes were distributed as shown according to the shares that define the prize pools: 0.22 for 4-ball, 0.10 for 5 ball, 0.16 for 5+bonus, and 0.52 for 6-ball.

Table 3 shows two suggested alternative ways of distributing the prize pool that has been suggested in Moore (1997). These represent an attempt to reduce the jackpot size by increasing the share of the 5+bonus pool (Scheme C), or increasing the share of the 5-ball pool (Scheme B), keeping the other prize pools constant. Both of these schemes would reduce the skewness of the prize distribution and this would be expected to reduce sales according to our estimates. However, they also reduce the variance and this, according to our estimates, should increase sales. The distribution of the number of winners remains the same since n and N that determine the game design is being kept at 6/49. Moore helpfully calculates the effect of this change in the

[23] A variable was also included which took account of the death of Diana, Princess of Wales. The television show was cancelled and many retail outlets were also closed. Finally, quarterly variables were included to represent any seasonal effects that might be present. For example, during the winter sales may be higher, as TV viewing goes up and more people watch the lottery show. Alternatively, during the summer, when there are fewer major news stories, the lottery may receive more media attention.

Table 3

The Allocation of Prize Money for the 62,476,486 Tickets Sold in the Week Ending 18.3.95

Prize types	Odds	Allocation of prize money	Expected no. of winners	Prize money allocated (£)	Actual no. of winners	Actual prizes awarded (£)
3 main numbers	1 in 57	£10 per ticket	1,096,131	12,778,130	1277,813	10
4 main numbers	1 in 1,033	22% of remainder	60,484	3,339,867	71,061	47
5 main numbers	1 in 55,492	10% of remainder	1126	1,533,070	1210	167
5 + bonus number	1 in 2,330,636	16% of remainder	27	2,454,000	25	98,160
All 6 main numbers	1 in 13,983,816	52% of remainder	4.5	7,975,572	2	3,987,786

Source: Moore (1997)

expected levels of prizes for each prize type and his calculations are reproduced in Table 4. The expected number of winners is calculated assuming that sales are £60m, a typical Saturday figure. Scheme *B* reduces the average jackpot win by £267,000 which allows the 4-ball prizes to be approximately £1,000 larger, while scheme *C* reduces the typical jackpot by more than £450,000 and this allows the typical 5+bonus winner to win more than £75,000 more.

The suggested alternatives have several effects on the characteristics of the prize distribution. First, as money is taken away from the large jackpot and moved to the smaller prizes, the mean increases. This is because the smaller prizes are extremely unlikely to roll over, so the total amount of money likely to be paid out on the current draw is higher. Thus, basis *A* gives the lowest mean and basis *C* the highest. Moving money away from the jackpot also decreases both the variance and the right skewness of the distribution. Compared to basis *A*, both *B* and *C* are less skewed and have smaller variance.

Table 4

Expected Monetary Prizes for Winners Under Three Schemes of Allocation

		Expected prizes (£) under the following schemes:		
Prize type	Expected no. of winners	Scheme *A* (52-16-10-22%)	Scheme *B* (45-16-17-22%)	Scheme *C* (40-28-10-22%)
Jackpot	4.3	1,988,749	1,721,032	1,529,807
5 + bonus number	25.7	101,987	101,987	178,477
5 numbers	1080	1,518	2,580	1,518
4 numbers	58,050	62	62	62
3 numbers	1,057,800	10	10	10

Source: Moore (1997)

Although it is harder to compare B to C, as C has a smaller jackpot but a larger 5+bonus prize than B, calculations of the moments of these distribution for plausible sales and rollover size imply C is less skewed and has smaller variance than B.

Decreasing the amount of money allocated to the jackpot prize also decreases the size of the rollovers, so that rollover draws under basis B and C will have a lower mean (offsetting the increase in mean for the regular draws) and have a smaller variance and skewness in comparison to rollover draws under basis A. In summary, for regular draws, a move from A to B to C increases the mean but decreases the variance and skewness, and decreases the value of all three moments for rollover draws.

To examine the overall effect on demand, sales for the first 200 lottery draws were simulated for each of the alternative prize allocations suggested above[24] compared to the simulated sales for the actual distribution of prizes. The results are presented in Table 5. According to the results of the simulation, basis C which is the least skewed, sells the most tickets. For non-rollover draws the sales increase from A through to C, implying the increase in mean and decrease in variance outweigh the effect of the decrease in skewness. Predictably, the size of the rollovers are smaller from A to C but the higher sales for regular draws also imply fewer rollovers under basis B and C than A. Interestingly, despite the smaller rollovers, average sales for rollover draws are still higher under basis B and C than under A. This is because the larger rollover and the large jackpot imply an increase in skewness and decrease in variance which is not outweighed by the increase in mean.

The implications are that the attractions of the lower variance and higher mean in the alternative schemes outweigh the detraction of their lower skewness: this generates higher sales in regular draws; this higher level of sales depresses the rollover probability so there are fewer rollovers; and while rollover sales are higher than sales in regular draws, this is not enough of a difference to outweigh the higher sales that occur under basis B and C.

Table 5
Simulation Summary

	Total sales over 200 draws (£m)	Total number rolled over (£m)	Number of rollovers	Average rollover size (£m)	Total rollover sales (£m)	Average rollover sales (£m)	Total non rollover sales (£m)	Average non rollover sales (£m)
Basis A	11,096	139.9	15	9.32	994	66.32	10,101	54.6
Basis B	11,912	85.8	10	8.58	709	70.92	11,203	59.0
Basis C	12,358	78.7	10	7.87	731	73.15	11,626	61.2

[24] Since the mean, variance and skewness depend on the level of sales, forecasting sales amounts to solving a highly complex nonlinear equation in sales. This proved hard to solve analytically and, instead, a recursive algorithm was used. There is, however, no guarantee that the solution will tend to a limit, or that any limit that does exist will be unique. In the examples shown here, the system converged very quickly and seemed invariant to the initial value of sales used.

Thus, despite what Camelot finds about what the public says that it wants, our evidence, based on what the public does when asked to dig into its collective pocket, suggests that it is quite possible to redesign the game to promote sales at the same time as reducing the typical jackpot size.

Attempts to simulate the sales for different takeout rates failed, as the iterations did not converge. Instead, the impact a change in takeout would have on sales was assessed by examining the slope of the sales function with respect to the takeout rate. This only captures the static effect of a change in takeout – the dynamic effects which are transmitted through rollovers are not accounted for.

Our results of this exercise are only suggestive: they suggest that an *increase* in the takeout rate would *decrease* sales but *increase* tax revenue. The higher frequency of rollovers caused by the lower sales would counteract this conclusion (but the rollovers would be smaller and so there would also be a loss of sales due to this). Thus, it would appear that the current game is 'too generous' – it would be worth increasing the takeout rate despite the reduction in sales that might ensue. However, we cannot say just how mean the game should be.

The final change in game design considered was the effect of changing the format of the game from 6/49 design currently used to a 6/53 design. The latter format was that proposed in the People's Lottery bid.

Our attempt to evaluate the possible impacts of this change once again did not converge.[25] Neither was it possible to examine how sales would vary with a small change in these design parameters because of their discrete nature. However, in an attempt to draw at least tentative conclusions, Table 5 compares the value of each of the moments for the two different game designs evaluated at typical values of sales. The final column gives the 'predicted level of sales' relative to the base of 6/49 with sales of £60m which is simply obtained by setting the values of the moments at the levels relevant to the assumed sales levels for each game design.[26] The moments are computed using the present arrangements for sharing the prize pool (we assume that there is also a bonus ball in the 6/53 design). Figs 2 and 3 suggested that, at least at high levels of sales, the effects of sales on the moments of the prize distribution is quite small – most of the variation in these variables arises from rollovers. Thus, in Table 5, we are assuming that the effect of sales on the moments is small enough to be ignored and we compute the predicted sales at the calculated levels of the moments corresponding to the chosen figures for sales and rollover size. We base the predictions at sales of £60m in a regular draw using the 6/49 design.

Increasing the number of balls in the draw makes all prizes more difficult to win and the mean for any draw therefore decreases. However, since fewer

[25] Several other attempts were made to resolve this, for example, by using the 'findroot' command in *Mathematica*, which solves non-polynomial expressions using the Jenkins–Traub algorithm. However, even with considerably simplified expressions for the moments, this also failed to produce a solution.

[26] Allowing for the typical Wednesday draw to be less popular than the typical regular Saturday draw by the estimated value

people are expected to win a share in each of the pari-mutuel prizes, those who do win can expect to win more. The variance and skewness therefore increase. These effects can be seen by comparing any row of panel (a) with the corresponding row in panel (b). The implications for sales are shown in the final column. Comparing, for example, 6/49 at a regular Saturday draw level of sales of £60m with the same draw under 6/53 the model predicts that sales would be lower by £11.5m. Comparing a regular Wednesday sales in 6/49, we find that sales would be £31.4m lower than the Saturday, but under 6/53 sales would be even lower than this – by a further £11.1m. Thus, for each of the points where the sales function was evaluated, the effect of the increased variance and decreased mean outweigh the effect of the increased skewness – sales would be about £21m lower per week under 6/53 compared to 6/49 *if* the shares and the take-out rate were kept fixed.

Against this lower level of sales in any particular draw has to be set the higher probability of a rollover. As Table 6 suggests, rollovers imply that sales would be higher: about £14m higher[27] on a Wednesday rollover draw and about £11.8m on a Saturday rollover. The hypothetical probability of a rollover for each game design can be computed assuming that individuals choose their numbers randomly. Note that these hypothetical numbers are considerably different from actual experience of the 6/49 game: the *theory* suggests that Saturday (Wednesday) rollovers should occur about once every 70 (9) draws; *in practice*, it is more like every 7 (3). Thus, these figures drastically underestimate the likely number of rollovers. However, they are probably a good guide to the *change* in the frequency of rollovers if we moved from 6/49 to 6/53.[28] Saturday rollovers should be about four times more frequent and

Table 6

Effect of Change in Game Format

Sales (£m)	Rollover (£m)	Mean	Variance	Skewness	Predicted sales relative to base
(a) 6/49 format					
60	0	0.4480	3.74×10^5	1.41×10^{12}	–
80	10	0.5741	1.26×10^6	8.50×10^{12}	11.8
30	0	0.4333	3.03×10^5	9.14×10^{13}	−31.4
40	10	0.6775	2.62×10^6	2.54×10^{13}	−17.3
(b) 6/53 format					
60	0	0.4382	6.94×10^5	4.34×10^{12}	−11.5
80	10	0.5662	2.39×10^6	2.94×10^{13}	−3.6
30	0	0.4065	4.07×10^5	1.58×10^{12}	−42.5
40	10	0.6298	3.46×10^6	4.36×10^{13}	−35.6

[27] That is, £31.4 – £17.3m.

[28] In fact, this comparison might even underestimate the rollover probability under 6/53 in practice because 53 has more numbers above 31 than does 49 so a higher proportion of the available combinations lie outside the range within which birthdays lie. Thus, 6/53 may experience a higher degree of conscious selection that does 6/49 and hence an even higher number of rollovers than we would expect.

Wednesday rollovers would be about twice as frequent. Since Wednesday sales are predicted to be about £14m higher in a rollover and Saturday sales are predicted to be about £12m higher in a rollover draw, weighting the predicted sales figures together by the rollover probabilities suggests that, on average, the 6/49 game would generate weekly (Wednesday plus Saturday) sales of around £95m while the 6/53 game would generate weekly sales of around £85m.

Thus, the higher rollover probabilities only compensate for about half of the loss in sales due to the change in the moments. Over the course of the licence, 6/53 might sacrifice £3.5b relative to 6/49. This is not to say that 6/53 *could* not be *more* successful than 6/49, but it would have to be combined with other changes: the takeout rate would need to be dropped a little to stop the mean return under 6/53 falling too far to depress sales in regular draws, and the higher variance in the prize distribution under 6/53 would need to be addressed, perhaps by dropping the bonus ball (and hence the 5+bonus prize pool) and adding it, instead, to the 3-ball prize pool.

Moreover, it would need to be marketed in such a way that sales in regular draws were encouraged: the danger with high rollover probability games is that the loss in sales that occurs in regular draws is not compensated by the occasional bout of lottomania that occurs when multiple rollovers have accumulated.[29]

However, our results need to be qualified. The simulations assume that sales respond to variations in mean, variance and skewness from design changes in the same way that they respond to these variables when rollovers occur. However, it is plausible that people may respond differently to these two types of changes. First, changes induced by occasional rollovers allow for the possibility of substitution between draws, but this possibility does not exist for changes coming through the game design rather than rollovers. This suggests that ticket sales are higher when changes come from rollovers than from game design. Rollovers are rather like sales promotions – they may induce people to change their behaviour quite differently to a temporary difference in the offer than they would to a permanent change. However, in the absence of a well-designed social experiment we cannot overcome this problem.

Second, our analysis is based on an econometric model of sales that, while it explains a high proportion of the variation in sales over time (as do many aggregate time series models), may not be good at forecasting the effects of structural changes (as is the case for many aggregate time series models).

[29] Sales may also be affected by the change in design because of behavioural considerations, which would not be picked up in our modelling. For example, getting two numbers right may lead the player to feel some measure of success and encourage him to play again, even though he won no prize. With a 51 or 53 board, the likelihood of getting 2 numbers is smaller and may leave the player feeling discouraged or bored. To give another example, a common pattern of play is to 'reinvest' small winnings, for example from getting 3 balls correct, in further play. Since the likelihood of getting 3 balls right is decreased with a 51 or 53 board game, this may again contribute to reduced sales. Moreover, the 3-ball prize pool not only serves to reduce variance, it also serves as an advertisement – with 1 in 57 winning a 3-ball prize under 6/49 players are likely to know someone who has won in the recent past. Under 6/53, there would be fewer 3-ball winners.

Third, the econometric model itself fails some of the diagnostic tests that were applied to it suggesting that it is misspecified is some way.

Finally, the model has not been validated by investigating how well it predicts structural breaks since none have occurred. Thus, one avenue for further research would be to apply the methodology to other places where design changes have occurred: Israel, Ireland and several US states spring to mind.

5. Summary and Conclusion

Our analysis has considered some of the most important questions relevant to running a lottery. Our methodology for analysing the implications of game design is, as far as we are aware, the most analytically rigorous yet to be applied to this issue, and yet it reflects the informal received wisdoms that dominate industry debate. Thus, it probably captures many of the important features of realities of the game but provides a degree of abstraction from reality to allow counterfactual changes to be analysed in a formal and quantitative way.

Our analysis is computationally very complex and relies on numerical methods to simulate the effects of reforms – there is no guarantee that the method will always produce a solution. Thus, our analyses of changes in the take out rate and in the format of the game are more speculative. They do *suggest* that the take out rate is too low – more revenue could be raised if the take-out rate were increased despite the drop in sales. They also suggest that 6/53, while it has some attractions, would, without other changes, lower the mean return to playing, raise the variance, as well as raise the skewness. The first two effects would lower sales and only the latter would raise them and, without other changes, the overall result would be lower sales.

Overall, our results support the U-turn made by the Lottery Commissioners: the analysis suggests that, *ceteris paribus*, sales would be lower with 6/53 than 6/49.

University of Warwick

National Economic Research Associates

Appendix: Calculation of Moments

Consider a lottery with a total of $p + q$ prizes, the first p of which are pari-mutuel prizes and the next q are fixed-value prizes. The ith prize is won with a probability π_i. The amount won for each of the fixed value prizes is given by W_i. For the ith pari-mutuel prize, the jackpot is given by J_i but the amount won depends on the number of other people who have correctly guessed the numbers. Since S is large, the number of winners can be approximated by a Poisson distribution so that the probability of j winners for prize i if S tickets are bought is given by: $P_{ij} = \mathrm{e}^{-S\pi_i}(S\pi_i)^j/j!$. The expected value of a lottery ticket can then be approximated by

$$\begin{aligned}\mu_1 &= \sum_i \sum_j \pi_i P_{ij} J_i/(j+1) + \sum_i \pi_i W_i \\ &= \sum_i \sum_j \pi_i [\mathrm{e}^{-S\pi_i}(S\pi_i)^j/j!] J_i/(j+1) + \sum_i \pi_i W_i \\ &= \sum_i \pi_i J_i \mathrm{e}^{-S\pi_i} \left[\sum_j (S\pi_i)^j/(j+1)!\right] + \sum_i \pi_i W_i \\ &= \sum_i \pi_i J_i \mathrm{e}^{-S\pi_i} \left[\sum_j (S\pi_i)^{j+1}/(j+1)!\right] \Big/ S\pi_i + \sum_i \pi_i W_i\end{aligned}$$

Approximating S by infinity, the second summation can be simplified by using the exponential series, so that

$$\mu_1 \cong \sum_i \pi_i J_i \mathrm{e}^{-S\pi_i}(\mathrm{e}^{S\pi_i} - 1)/S\pi_i + \sum_i \pi_i W_i = \sum_i J_i (1 - \mathrm{e}^{-S\pi_i})/S + \sum_i \pi_i W_i$$

The second and third moments are given by

$$\begin{aligned}\mu_2 &= \sum_i \sum_j \pi_i P_{ij} [J_i/(j+1)]^2 + \sum_i \pi_i W_i^2 \\ &= \sum_i \sum_j \pi_i [\mathrm{e}^{-S\pi_i}(S\pi_i)^j/j!][J_i/(j+1)]^2 + \sum_i \pi_i W_i^2 \\ &= \sum_i \pi_i J_i^2 \mathrm{e}^{-S\pi_i} \sum_j (S\pi_i)^j/[(j+1)(j+1)!] + \sum_i \pi_i W_i^2\end{aligned}$$

$$\begin{aligned}\mu_3 &= \sum_i \sum_j \pi_i P_{ij} [J_i/(j+1)]^3 + \sum_i \pi_i W_i^3 \\ &= \sum_i \sum_j \pi_i [\mathrm{e}^{-S\pi_i}(S\pi_i)^j/j!][J_i/(j+1)]^3 + \sum_i \pi_i W_i^3 \\ &= \sum_i \pi_i J_i^3 \mathrm{e}^{-S\pi_i} \sum_j (S\pi_i)^j/[(j+1)^2(j+1)!] + \sum_i \pi_i W_i^3\end{aligned}$$

These last two series can be shown to converge by using a simple ratio test which states: given an infinite series of positive terms $a_1 + a_2 + a_3 + a_4 + \cdots + a_n + a_{n+1} + \cdots$ if the ratio a_{n+1}/a_n tends to a limit A as n tends to infinity then if $A < 1$ the series converges (absolutely) and when $A > 1$ the series diverges. If $A = 1$, the test gives no information. The variance and skewness (central moments) can then be derived using the expansions below although, in practice, these barely differ from the moments because the mean is so small (of order 1) compared to the second moment and third moments, which were of order 10^5 and 10^{12} respectively: $(X - \mu_1)^2 = \mu_2^2 - \mu_1^2$ and $\mathrm{E}(X - \mu_1)^3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3$.

The size of the jackpots were calculated by approximating the amount paid out in three ball prizes as $S\pi_3$ and then dividing the remaining prize money as laid out in section 3. The moments were then evaluated by truncating each of the infinite series to the first fifty terms, which ensured the calculations were accurate to five significant figures. To ensure numerical precision during estimation, the second moment was scaled by 10^5 and the third moment by 10^{12}.

References

Beenstock, M., Goldin, E. and Haitovsky, Y. (1999). 'What jackpot? The optimal lottery tax', The Hebrew University of Jerusalem, mimeo.

Clotfelter, C. T. and Cook, P. J. (1991). *Selling Hope: State Lotteries in America*, New York: Harvard University Press.

Cook, P. J. and Clotfelter, C. T. (1993). 'The peculiar scale economies of Lotto', *American Economic Review*, vol. 83, pp. 634–43.

Farrell L., Lanot G., Hartley R. and Walker, I. (2000). 'The demand for lotto and the role of rollovers and conscious selection', *Journal of Business and Economic Statistics*, vol. 18, pp. 228–41.

Farrell, L. Morgenroth, E. and Walker, I. (1999). 'A time-series analysis of UK lottery sales: long and short run price elasticities', *Oxford Bulletin of Economics and Statistics*, vol. 61, pp. 513–26.

Farrell, L. and Walker, I. (1999). 'The welfare effects of lotto: evidence from the UK', *Journal of Public Economics*, vol. 72, pp. 99–120.

Forrest D., Gulley, D. and Simmons, R. (2001). 'Elasticity of demand for UK National Lottery tickets', *National Tax Journal*, vol. 53, pp. 853–63.

Forrest, D., Simmons, R. and Chesters, N. (2001). 'Buying a dream: alternative models of demand for lotto', Centre for Gambling and Commercial Gaming, University of Salford, Discussion Paper 01–01.

Garrett, T. A and Sobel, R.S. (1999). 'Gamblers favor skewness, not risk: further evidence from United States' lottery games', *Economics Letters*, vol. 63, pp. 85–90.

Golec, J and Tamarkin, M. (1998). 'Bettors love skewness, not risk, at the horse track', *Journal of Political Economy*, vol. 106, pp 205–25.

Moore, P. G. (1997). 'The development of the National Lottery: 1992–96', *Journal of the Royal Statistical Society (Series A)*, vol. 160, pp. 169–85.

Munting, R. (1994). *An Economic and Social History of Gambling in Britain and the USA*, London: Manchester University Press.

Scoggins. J. F. (1995). 'The Lotto and expected net revenue', *National Tax Journal*, vol. 48, pp. 61–70.