Modelling demand for lotto using a novel method of correcting for endogeneity*

by

Rose Baker University of Salford, UK

David Forrest University of Liverpool, UK

and

Levi Pérez[†] Universidad de Oviedo, Spain

April, 2019

* We acknowledge the financial support of *Loterías y Apuestas del Estado*. Pérez acknowledges financial support from grant FC-15-GRUPIN14-064 of the Government of Asturias, Regional Council of Economy and Employment. We are grateful to The Editor and two reviewers whose comments and suggestions have been of considerable assistance in improving the paper. We also thank Prof. Ian McHale for fruitful discussion during the course of the project.

† corresponding author

Abstract: Modelling lottery sales as a function of the mean, standard deviation and skewness of the probability distribution of returns potentially gives insights into how the design of a game could be modified to maximise net revenue. But use of OLS is problematic because the level of sales itself affects values of the moments (and insufficient instruments are available for IV regression). We draw on the concept of a rational expectations equilibrium, developing a new regression model which corrects for endogeneity where the causal impact of the dependent variable on the right-hand side variables is deterministic. Results provide more reliable guidance to lottery agencies because accounting for endogeneity leads to significantly different results from OLS and these results have superior performance in out-of-sample forecasting of sales. More generally, results prove consistent with the Friedman-Savage explanation of why people buy lottery tickets and with evidence from racetrack data that 'bettors love skewness'.

Key words: lottery; risk preferences; rational expectations; endogeneity; regression
JEL codes: C50, H27, L83

Highlights:

- **we model lotto demand, correcting for endogeneity without the use of instruments**
- **assuming rational expectations allows removal of biases caused by reverse causation**
- **our new method for estimating lotto demand provides improved forecasting of sales**
- **lotto players respond positively to expected value and skewness**

1. Introduction

A lotto game offer players the chance to win a large jackpot prize. The player pays an entry fee and chooses numbers from a set of numbers according to the format of the particular game. For example, players may be asked to choose any six numbers from the set 1-49 (or have the computer choose for them). After the draw closes, a proportion of revenue from entry fees is paid into a jackpot pool. A random draw of numbers then takes place, for example six balls are drawn from a container with 49 numbered balls. These are the winning numbers. Any player whose number selection exactly matches the numbers drawn wins a share of the jackpot pool (the sales revenue also funds smaller prizes for ‘near misses’, for example for entries which match only four or five of the six balls). In the event that no one wins a share in the jackpot, the money allocated ‘rolls over’ to the following draw, i.e. it is added to the jackpot prize pool for that draw. This will make the game better value next time it is offered.

Economists have long been interested in modelling demand for such lotto games, for two main reasons. First, they are typically state-operated or state-sanctioned and generate considerable revenues, which are often dedicated to funding ‘good causes’. Second, economists have been intrigued by the popularity of lotto games given the typically poor returns-to-player on offer. Capturing consumer preferences informs choices on how game design may be modified to maximise revenue for good causes. Potentially, studying demand will also offer insight on why people buy lottery tickets in the first place and add to understanding of how risk preferences can explain the phenomenon of gambling.

Researchers modelling lotto demand (for a detailed survey, see Pérez and Humphreys, 2013) rely on the ‘rollover’ feature of the game for identification of the demand function. Whenever the jackpot prize is not won, the jackpot money carried forward to the following draw raises the size of the jackpot even if some of the smaller prizes remain the same. Often several consecutive rollovers occur and the jackpot prize becomes many times greater than in the first draw in the sequence. So draws observed over any given period may differ very considerably from each other in terms of the set of prizes available to the purchaser of a ticket. It is this variation from draw to draw which facilitates identification.

Essentially the purchaser of a ticket is regarded in this research as buying a probability distribution of winnings, which includes a high probability of winning nothing, a low probability of winning a modest amount, and a very low probability of winning an extremely high prize. Rollovers change the characteristics of this probability distribution of winnings. For example, they increase the mean (or expected value) because they add money from preceding draws to the prize pool for the current draw; and they also change the shape of the distribution, because all the additional funds carried forward from the preceding draw are allocated just to the top prize, raising skewness. Relating the sales in each draw to the characteristics of the probability distribution of prizes generates a demand function defined in terms of those characteristics.

The characteristics of the probability distribution of winnings in a draw are conveniently summarised in terms of mean, variance (or standard deviation) and skewness (as in, for example, Walker and Young, 2001). But the researcher runs into a problem in attempting to estimate demand as a function of these moments. Purchasers can indeed be expected to respond to variations in mean, variance (or standard deviation) and skewness. However, there is also reverse causation. Fixed proportions of sales revenue from the current draw are paid into the jackpot and lower prize pools, which affects the values of each of these moments.

This and other types of endogeneity in economic modelling which lead to biased coefficient estimates are classically resolved by resort to the use of instrumental variables. Gulley and Scott (1993) estimated lotto sales as a function of expected value and proposed that expected value could be instrumented by the amount rolled over from the preceding draw. Subsequent authors (such as Farrell et al., 1999) followed their example.

However, specifying sales as a function of just expected value is an incomplete representation of consumer preferences. Expected value is less than the price of a ticket, so the decision to buy appears likely to be driven by higher moments of the probability distribution of winnings. Unfortunately, a more complete specification would require additional instruments (one for each of the additional variables) and none have been identified. Thus, Walker and Young (2001) had to estimate UK Lotto sales as a function of mean, variance and skewness by ordinary least squares (OLS) even though all three variables were measured ex post and were therefore endogenous. How much this will have biased coefficient estimates has remained unclear.

Our methodological contribution is to show that the problem of endogeneity can in fact be resolved when estimating lotto demand without the use of *any* instruments. This is only possible because reverse causation in this particular case is deterministic: the causal impact of the dependent variable on the explanatory variables is determined by the operator's fixed formula for allocating a share of revenue from ticket sales in a draw to prizes for that draw. In this context, a demand curve purged of the biases associated with endogeneity may then be estimated by drawing on the concept of a rational expectations equilibrium. In the rational expectations equilibrium, where the underlying assumption is that agents on average make unbiased forecasts, predicted sales used to compute the values of the moments are required to equal the sales predicted by the moments model. Our methodology extends the scope of rational expectations theory to a new market setting. In the past, it has been used primarily in macro-oriented contexts but also underpins much analysis in the area of asset prices (Sargent, 2008).

We apply our rational expectations approach to a data set of sales for the principal lotto game offered in Spain. We demonstrate that estimation results are materially different from OLS results. Coefficient estimates are smaller than those obtained from OLS because the model has been purged of endogeneity. This is practically important because reliably estimating consumer preferences over mean, variance and skewness allows operators to predict consumer response were the pattern of returns to be altered by changing the format of the game.

Finally, by offering a reliable way of using lotto sales models to reveal consumer preferences over the characteristics of the probability distribution of returns, we will be able to confirm whether the pattern of signs on mean, standard deviation and skewness is consistent with the Friedman-Savage utility-of-wealth function (Friedman and Savage, 1948). Confirmation of the relevance of Friedman-Savage from experimental data has been hard to obtain because experiments cannot realistically offer sufficiently large gains which reveal preferences for gambling for life-changing prospective gains. But Brunck (1981) provided indirect evidence from a National (US) survey where respondents' dissatisfaction with current income predicted participation in lotteries but not in other gambling activities where extreme wins are not offered. And, over time, study of gambling markets has added slowly to evidence through analysis of naturalistic data. Golec and Tamarkin (1998) found patterns of odds and returns across horses that were consistent with racetrack bettors being risk-averse but skewness-loving, which would be the case if they had Friedman-Savage utility functions. Walker and Young (2001) took a yet more direct approach by regressing lottery draw sales on the mean, variance and skewness of returns in the particular draw and found coefficient estimates to be positive, negative and positive respectively (albeit that skewness was only marginally significant). This was evidence consistent with Friedman-Savage; but, as noted already, they had to use OLS despite acknowledging the presence of endogeneity. By resolving the endogeneity issue, we are able to consider safer estimates of the relevant coefficient estimates and test whether, for lottery play at least, behaviour in the field is consistent with the Friedman-Savage assumptions and with their rationale for lottery play: purchasers of lottery tickets are sufficiently compensated by high positive skewness in returns that they buy tickets which are poor value in terms of the mean return.

The remainder of the paper is structured as follows. In Section 2, we provide context for the Spanish data set we analyse. We go on to work with a model where sales are a function of the moments of the probability distribution of winnings. Section 3 outlines issues on how these moments may be measured. In Section 4, we build our model, based on the idea of a rational expectations equilibrium, and seek to justify our claim to have resolved the roadblock in modelling lotto demand caused by the lack of obvious instruments. Our strategy for computing the parameter estimates and standard errors is described in an Appendix. The estimates themselves are presented in Section 5 and compared with OLS estimates. We show that estimation from our modelling based on rational expectations provides more accurate out-of-sample sales forecasts than OLS and that the material difference between coefficient estimates is large enough to be relevant in practical application by those evaluating prospective changes in the structure of any lotto game. The final section of the paper reflects on what has been learned.

2. Data and context

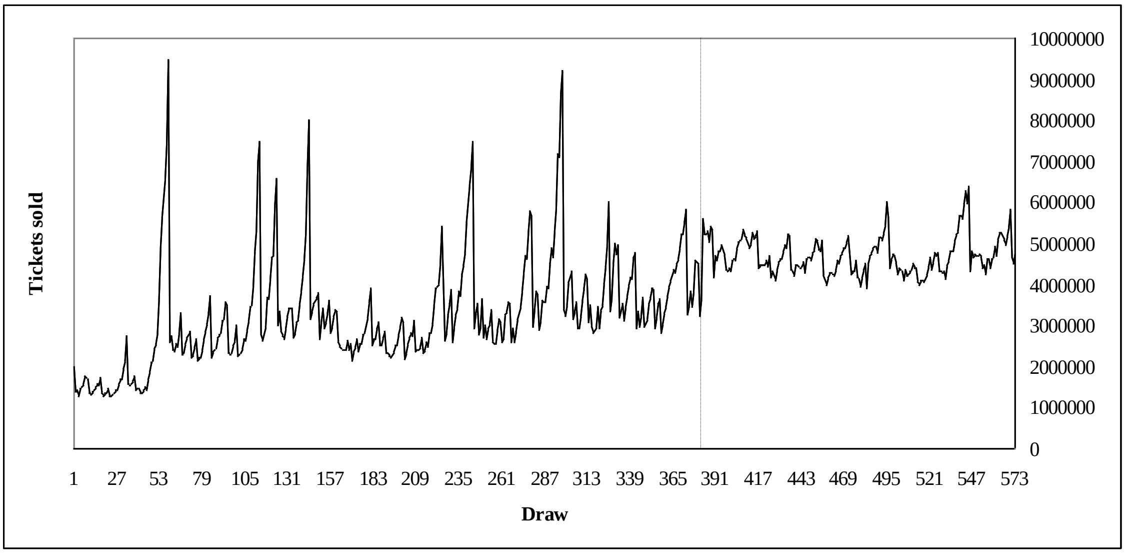
The game we study to illustrate our methodology is *El Gordo de la Primitiva*, offered by the Spanish state lottery agency, Loterías y Apuestas del Estado (LAE). It was launched as a monthly game in 1993 but converted to weekly play on October 12, 1997, and the data to be employed in our analysis relate to draws from then on (and up to September 28, 2008). This gave us 573 weekly observations. Throughout these weeks, the cost of purchasing a ticket (€1.50) and the proportion of sales revenue earmarked for prizes (55%) remained the same.

Initially the game was sold with the single matrix m/n format used in many other jurisdictions, with m=6 and n=49. So players had to choose six numbers from the set 1 to 49 and, if the selection made exactly matched the six main numbers drawn (the probability of such a match is approximately 1 in 14m), the player shared in the jackpot prize. There were also four lower prize tiers and the additional possibility of a refund of the entry fee, which was made to holders of 10% of the tickets sold (chosen by a separate random process). Detailed rules were in place for the proportion of ticket revenue allocated to each prize level.

On February 6, 2005, at the 383rd draw in our data set, LAE introduced a major modification to the design of the game by changing to a two-matrix format, 5/54 + 1/10. So now a player had to choose five numbers from the set 1 to 54 plus an additional number from a second matrix consisting of the ten numbers from 0 to 9. Consequently the chance of winning a share in the jackpot fell to 1 in nearly 32m. The entire prize structure was changed, with eight instead of five tiers of prize (in addition to the refund, for which the probability remained 0.1). Table 1 presents a summary of the basic rules of the game before and after the changes in game design. Figure 1 shows draw-by-draw sales figures

It should be noted that, before each draw, LAE announces how much has been rolled over from preceding draws but does not issue forecasts of projected jackpot. Jackpot size is known only after sales close because a proportion of sales revenue is added to any rolled over funds already in the pool. Players' decisions on how many tickets to buy are therefore made without complete information and must be based on *expectations* concerning the distribution of prizes on offer.

Figure 1. Numbers of tickets sold for *El Gordo de la Primitiva*



Because the game was made harder to win by the design change in 2005, the jackpot was won less often, despite an increase in sales, and draw cycles were therefore typically longer. This allowed jackpots to accumulate to larger amounts than had been observed under the old format and the highest recorded in our data was €26.7m. Through the sequence of draws before the jackpot was won, the patterns of expected value, standard deviation and skewness were radically different from before. High skewness was more commonly a feature than under the previous arrangements. How sales evolved is shown in Figure 1 where the vertical line shows the point at which the design change came into effect.

Table 1. Rules and prize structure of *El Gordo de la Primitiva*

	Before February 6, 2005			After February 6, 2006		
Format	6 from 49			5 / 54 + 1 / 10		
Drawing frequency	Weekly			Weekly		
Ticket price (€)	1.5			1.5		
Take-out rate[a]	0.45			0.45		
Prize categories	5			8		
Share of the prize pool[b], number of balls to be matched and probability of winning						
Jackpot	0.55[c]	6	7.151×10^{-8}	0.22[d]	5+1	3.162×10^{-8}
2nd category	0.05	5+1[e]	4.291×10^{-7}	0.33	5+0	2.846×10^{-7}
3rd category	0.16	5	1.845×10^{-5}	0.06	4+1	7.747×10^{-6}
4th category	0.24	4	9.69×10^{-4}	0.07	4+0	6.972×10^{-5}
5th category	15.03	3	0.0177	0.08	3+1	3.72×10^{-4}
6th category	-	-	-	0.26	3+0	0.00335
7th category	-	-	-	0.2	2+1	0.00583
8th category	-	-	-	3	2+0	0.0524

Notes: [a] take-out rate is the proportion of entry fees retained by the operator to cover operating costs and profit. [b] 55% of total income goes into the prize pool, but 10% goes to a fund for the refund of the ticket price prize and the remaining 45% is then distributed among prize 'categories'. [c] Once the total amount devoted to the fixed prize for the 5th category has been deducted from 45% of total income, the remaining amount is distributed among prize categories (including the jackpot). [d] 22% of total income goes directly into the jackpot prize pool. The remaining 23% of total income – after deducting the total amount devoted to the flat prize for the 8th category – is distributed among lower categories. [e] A seventh ball was drawn before February 6, 2005. Matching 5 numbers and the 'Bonus Ball' won the second highest prize.

We note that the design change in question was not a response to falling sales (see Figure 1). Change therefore appears to have been a genuine experiment to confirm the agency's view on how players would be likely to respond to return-risk-skewness packages that were different from those that had been available through the draw cycle in the past. It provides for our empirical analysis something akin to a natural experiment where change can be considered to be exogenous (until recently relatively few other jurisdictions had altered the design of their lotto games and those that did so were typically responding to falling sales).

3. Measurement of the moments

Given that we intend to model lotto sales as a function of expected value, standard deviation and skewness, it should be noted before we begin that there is some ambiguity in how values for the three moments should be calculated. Previous papers calculate the values for a ticket for which the combination of numbers has been randomly selected and on the assumption that all other entries to the draw are also based on random selection. However, many players choose numbers for themselves rather than opt for a computer generated entry. These choices are liable to be correlated with each other and any given level of sales will be associated with a lower proportion of the possible number combinations being selected. This is the issue of 'conscious selection', raised first by Cook and Clotfelter (1993). Conscious selection lowers the probability that the jackpot will be won at all and produces increased variability in the number of winners who share the jackpot when it is won.

To take account of conscious selection when calculating the moments, we adapted methodology developed in Baker and McHale (2009). Our main results will be from modelling where the moments have been calculated following the Baker-McHale approach. However, we checked the sensitivity of results to the assumption of conscious selection and the methodology for allowing for it by repeating our estimation but with moments calculated *without* allowance for conscious selection (i.e. every player was assumed to pick his or her numbers randomly). Results were in fact very similar, validating and extending the claim in Farrell et al. (1999) that allowing for conscious selection when modelling UK lotto sales as a function of expected value, made no material difference to results. It had not been clear to us whether the same would apply when higher moments were added to the specification but, in the event, we drew the same conclusion. Therefore our final results will not be dependent on decisions made about dealing with conscious selection.

4. A model for resolving endogeneity

4.1 Overview

As the moments used to predict lotto sales are themselves a function of sales, OLS must be modified in some way because of endogeneity; but, as we have noted, insufficient instruments are available for IV regression to be an option. Here we adopt instead a method based on constructing a self-consistent estimate of sales from the regression equation. This self-consistent estimate is not a function of observed sales, q, hence removing the endogeneity problem.

This procedure has a simple economic interpretation. There are many econometric models in which people's expectations figure as predictive variables (Murphy and Topel, 2002). Greene (2008, p. 507) gives the illustrative example of a model in which the expected number of children could be a predictor variable in the decision to enrol in job training. In our case, the difference is that expected sales influences the values of the moments of the prize probability distribution which, in turn, influence the prediction of sales from the regression equation. This leads to a self-consistency condition that must be met in a rational expectations equilibrium. One can imagine the potential purchaser of a ticket hypothesising a likely sales figure, evaluating the resulting attractiveness of the lottery, and hence refining his or her estimate of sales until a self-consistent estimate is reached.

As this methodology, which we term self-consistent regression (SCR), is new, computations were done by writing fortran95 programs. These used Numerical Algorithms Group (NAG) routines for random number generation, function minimisation, numerical differentiation, and matrix inversion. Full description of our computation strategy is provided in an Appendix. In the remainder of the present section, we first describe the model itself and then explain why and how it allows unbiased estimates to be derived despite endogeneity issues and despite the lack of appropriate instrumental variables.

4.2 Model description

The observed sales figure is modelled as

$$q_t = \sum_{j=1}^{p} \beta_j x_{jt}(\hat{q}_t) + \eta_t \qquad (1)$$

where the subscript t signifies the draw number. The β_j are regression coefficients and the x_j are variables including expected ticket value, standard deviation and skewness (x_1=1 to permit β_1 to serve as an intercept). The model is specified as linear as the simplest choice where theory provides no guidance as to functional form.

Control variables include draw number (to account for any trend in sales), shift and slope dummies to account for any effects (other than through changes in the prize probability distribution) from the introduction of a new game design, and two lagged values of sales, to account for habit formation. Some of the x_j, specifically expected value, standard deviation and skewness, are functions of predicted sales $\hat{q}$. Finally, η is an exogenous random error about which we shall give further detail later.

Since by definition $q = \hat{q} + \eta$, we have the self-consistency condition

$$\hat{q}_t = \sum_{j=1}^{p} \beta_j x_{jt} \left(\hat{q}_t\right) \qquad (2)$$

This means that the forecast of sales that we imagine the player makes is also our best forecast from the trend and moments of the prize distribution. The only parameters in the model are $\beta_1,\ldots,\beta_p$, but the model is now nonlinear in these β_j because the predictor $\hat{q}$ solves a nonlinear equation.

4.3 The role of rational expectations

Our SCR model is essentially based on the notion of a rational expectations equilibrium, a concept applied generally to markets where participants must forecast the future or some other unknown: such a market is said to be in equilibrium when the expectations of market participants match actual outcomes on average (the wisdom of crowds). Here, the values of the moments are not known at the time of ticket purchase as they depend on finally realised sales. In the absence of published forecasts on the final size of the jackpot, and using (readily) available information such as the size of rollover paid into the jackpot pool for the current draw, players are assumed to make their own forecasts of sales in order to assess the values of the moments (they then decide how many tickets to purchase). The assumption does not imply that their forecasts are perfect but it does imply that the forecasts are unbiased and that any forecast errors are not serially correlated.

This of course is a strong assumption which may not necessarily hold. However, such evidence as there is from the prior literature on the rationality of lotto players suggests that it may be regarded as a fair working assumption. Scott and Gulley (1995) developed a method of testing for a rational expectations equilibrium in a lotto market. They found that actual sales were uncorrelated with the residuals from the fitted expected value equation in a two stage procedure, concluding that players, on average, correctly forecast sales using available information and supporting the idea of a rational expectations equilibrium. Forrest et al (2000) replicated this finding for the case of the market for UK Lotto and Matheson and Grote (2003) found, when looking at American lotteries, that players acted as if they were able to adjust their expected value forecasts in the face of new arrangements for paying out the jackpot prize as an annuity that appeared almost to constitute misinformation. This again is strong evidence of efficient use of data on the part of players and adds to the plausibility of assuming a rational expectations equilibrium in a lotto market.

How does assuming a rational expectations equilibrium enable an endogeneity issue to be resolved without recourse to conventional IV estimation?

First, one could state that the forecast $\hat{q}$ can be interpreted as akin to an instrumental variable, because it correlates with the predictor variables but not with the error term η. We need only one such variable, despite having several predictors, because they are all functions of the one quantity $\hat{q}$. The error η represents the influence of variables external to the lottery, such as the current weather or economic conditions, because we aim to have used all the lottery-based information such as rollover in our regression model. Hence $\hat{q}$ correlates with the x_j but not with η.

However, a deeper insight is that it is possible to remove endogeneity because we know sufficient information about the mechanism through which it operates to account for it in another way. A key point is that SCR deals only with one specific type of endogeneity. By contrast, the IV method is a very general tool that can be applied to all types of endogeneity, at the cost of introducing fresh variables into the analysis.

Here sales q are a function of predictors x, and vice versa. The functional dependence of q on x is specified by the model, but here the reverse functional dependence $x(q)$ is known. The regression need only be corrected for this latter dependence to make it work. How easy this is can be seen by a first-order (approximate) solution; expand $x(\hat{q}) \cong x(q) + x(q)'(\hat{q} - q)$, where $x'(q)$ is the (calculable) derivative dx/dq evaluated at observed sales q. Then a regression

$$q = \beta x(\hat{q}) + \eta \tag{3}$$

becomes

$$q \cong \beta x(q) + \{1 - \beta x'(q)\}\eta \tag{4}$$

Here, x is now evaluated only at observed sales q, and the only departure from OLS is that the random error η must be corrected by a function of β. This makes the regression nonlinear, so that the *approximate* solution could be found by nonlinear least squares (NLLS).

5. Results and model validation

Summary statistics are shown in Table 2 and results from estimation are displayed in Table 3.

The model is

$$q_t = \sum_{j=1}^{p} \beta_j x_j(\hat{q}_t) + \eta_t \tag{5}$$

where the sales figure for draw t (in millions) is modelled as a linear regression on p predictor variables, x_j, some of which (expected value, standard deviation and skewness) are themselves functions of sales, q_t, and others, such as lagged and doubly-lagged sales (included to capture habit formation) are not. We also include a trend term (draw number) and an interaction term where we multiply trend and a dummy variable (set equal to one for draws from Draw 383 on). This slope dummy permits trend to be different after the design change in *El Gordo de la Primitiva*.

Table 2. Summary statistics

	mean	standard deviation	minimum	maximum
sales (millions)	3.752	1.293	1.272	9.478
expected value	0.799	0.170	0.607	1.493
standard deviation	1098.5	915.3	78.0	4384.7
skewness	4529.1	901.5	3409.4	5903.1

Table 3. Regression results

	(1)		(2)		(3)	
		|t|		|t|		|t|
estimator	OLS		SCR		SCR	
conscious selection	yes		yes		no	
lag 1 sales	0.2851	10.42	0.2640	10.66	0.2675	10.80
lag 2 sales	0.0882	3.67	0.1013	4.04	0.1007	4.39
draw number	0.0014	6.63	0.0028	9.73	0.0028	9.40
draw number*new design	-0.0023	4.23	-0.0026	2.91	-0.0027	3.02
expected value	10.1456	21.52	6.493	12.09	6.435	8.98
standard deviation	-0.0015	14.41	-0.0007	5.34	-0.0007	4.50
skewness	0.0010	17.01	0.0004	7.81	0.0005	6.88
constant	-9.0036	19.28	-4.870	9.38	-4.987	7.40
observations	573		573		573	
R-sq	.913		.921		.922	

Note: p-value to three decimal places was < .001 for all coefficient estimates
OLS refers to ordinary least squares and SCR to self-consistent regression

The OLS results (where values of the moments were computed taking conscious selection into account) appear in the leftmost columns of Table 3. Here, the moments were computed using the realised sales in each draw that are to be predicted; this contradiction is the weakness of the OLS model.

Although the problem of endogeneity with an OLS specification appears obvious, we nevertheless carried out a Hausman specification test to confirm that the OLS estimator indeed suffered from endogeneity. The test was carried out for a subset of variables, i.e. the expected value, standard deviation and skewness coefficients, and gave the chi-squared statistic as χ^2 [3] = 200.73, showing that the OLS estimator is not consistent given the endogeneity problem it incurs.

The second set of estimates in Table 3 reports the corresponding result from the SCR method. The third column is for the case where we re-estimated the SCR model with no allowance for conscious selection. The similarity of the results between columns (2) and (3) demonstrates that whether or not conscious selection is allowed for when modelling lotto sales is not in fact an important issue.

The principal feature of Table 3 is that there is a material difference between the OLS and the SCR results. When estimation by OLS (column 1) is replaced by estimation with SCR (column 2), the coefficient on expected value falls to about 64% of its previous value and the coefficient estimates on standard deviation and skewness fall, respectively, to 47% and 40% of their previous values. This is to be expected given that the equation has effectively been purged of the effects of endogeneity. All coefficient estimates, however, remain very highly significant and of the expected signs.

The self-consistent model appears to track sales more accurately than OLS. Holding back the last 50 draws, the MAPE (mean absolute percentage error) of the out-of-sample forecast of sales from the SCR model (with conscious selection) for the last 50 draws was 4.85%. When we estimated the OLS model holding back the last fifty draws, and “predicted” sales for those fifty using values for the moments calculated according to realised sales, the MAPE was 9.80%, more than twice as high. The same story held if we made comparisons using the SMAPE (symmetric mean absolute percentage error) (4.93% versus 10.47%) or the MAAPE (mean arctangent absolute percentage error, see Kim and Kim, 2016) (4.84% versus 9.76%). This illustrates the potential gain from adopting the self-consistent model for practical purposes.

A footnote to our results is that the redesign of the game appears, in fact, to have been very successful in terms of total sales, which were 44% higher in the twelve months following than in the twelve months preceding the change of format. Given that inclusion of the moments as regressors already accounts for restructuring of the prizes on offer, the results on the trend variables suggest that players collectively found the game format itself less satisfactory than before (perhaps, for example, they found it irksome to select their numbers from two matrices rather than one). Therefore it is likely that the strong positive response of sales to the reform was indeed linked to their preferences regarding the prize probability distributions available over the twelve months before and after the design change. That the game became harder to win made draw cycles longer and there were therefore more weeks with high skewness. Further, at any given point in the draw cycle, skewness was higher than it would have been under the old rules. Given the preference for skewness revealed by the regression results, it appears reasonable to link increased demand to increased skewness.

6. Concluding remarks

Previous work attempting to model lotto demand encountered an endogeneity problem which was not resolved because of a shortage of appropriate instruments. We have attempted a resolution by developing and employing a new class of regression model which corrects for endogeneity in the special context where the causal impact of the left hand side variable on the right hand side variables is deterministic. Results were markedly different from using OLS as the best alternative in the absence of instruments. The coefficient estimates on expected value, standard deviation and skewness all fell substantially but in different proportions to each other, indicating that ignoring endogeneity may generate estimates that are misleading for operators seeking guidance on consumer preferences over aspects of game design. Despite its computational complexity, we therefore recommend use of our self-consistent regression given that the lotto industry is important in terms of both its scale and the social importance of the expenditures it funds.

Using results from modelling sales is of practical importance to the lottery industry. Game managers must decide on game formats (e.g. whether a game should be choose-six-numbers-from-49 or choose-7-from-51) and prize structures (e.g. how much of the prize pool should go to the jackpot?) In assessing alternatives, a key input is formal evaluation of how consumers respond to different packages of expected value, standard deviation and skewness in returns. Our methodology allows more reliable input into decision-taking.

A more general contribution of our paper is that we have identified a method which, without requiring instrumental variables, enables correction for endogeneity associated with reverse causation in the particular case where the influence of the dependent on the explanatory variables operates mechanistically. Whether our innovation in methodology can be applied to settings other than the lottery we leave open to future research. Precise knowledge of how reverse causation operates may in fact be a rare situation but we speculate that it may be present in the setting of network goods where the utility of a service depends on how many other people subscribe. For example, suppose subscriptions to an online poker room are modelled as dependent on the expected waiting time for a playing partner to be found. Here there is reverse causation which is likely to operate mechanistically, permitting the application of SCR where there is no instrument available for expected waiting time.

Our paper has also yielded findings which test the validity of the Friedman-Savage (1948) utility-of-wealth function. The positive/ negative/positive coefficient estimates on expected value, standard deviation and skewness (all strongly significant) are consistent with the shape of function they proposed and consequently with their explanation of why people buy lottery tickets. Previous attempts at validation of Friedman-Savage from naturalistic data on lottery sales may have been unreliable to the extent that endogeneity biased the coefficient estimates. Our methodology for removing endogeneity allows more confidence than before in the Friedman-Savage representation of risk preferences.

7. Appendix

This appendix provides details of the exercise in computation that was required to estimate our SCR model.

To compute parameter estimates and standard errors, a workable strategy is to evaluate the skewness and other moments, at a number of equally-spaced values of q, for each draw. In this study, forty values from 0.25 million to 10 million sales were used. The value of $x_j(\hat{q})$ is then found at any value of $\hat{q}$ by interpolation. Quadratic inverse interpolation was used to solve the equation

$$\hat{q}_t - \sum_{j=1}^{p} \beta_j x_{jt}(\hat{q}_t) = 0 \qquad (6)$$

What happens computationally is that the log-likelihood function is maximised using a function maximiser. At any stage en route to the optimum values β^*_j, the maximiser requires the log-likelihood to be evaluated at values of β_j of its choice. Then, for each draw, equation (3) is solved to yield $\hat{q}$, the predicted sales for that draw, and this value is used to compute the log-likelihood as shown below. The values of $x_j(\hat{q})$ that have been tabulated are specific to the particular draw, because they are functions of the amount rolled over and these values are used to compute the log-likelihood across all draws as shown below.

Concerning the error structure, it is reasonable to suppose that the error will likely be a percentage or proportion of sales. For example, on a sunny day, maybe 10% more people play, or those who do increase their purchases by 10%. This thought leads to a lognormal distribution for sales, $\ln(q_t) = \ln(\hat{q}_t) + \varepsilon_t$, where □$_t$~N[0, □^{2}]. In terms of the error □, we have □ = {exp (□) - 1}$\hat{q}$.

We tested the lognormal assumption by using the Box-Cox transformation to replace $\ln(q) = \ln(\hat{q}) + \varepsilon$ by

$$T(q) = \frac{q^{\lambda} - 1}{\lambda} \qquad (7)$$

The best fit gave the power as □ = 0.056, very close to zero, showing that a logarithmic transformation does give the best fitting model; the fit residuals are then approximately normally distributed.

The log-likelihood function is then

$$\ell(\beta) = -(1/2)\sum_{t=1}^{n}(\ln(q_t) - \ln(\hat{q}_t))^2 / \sigma^2 - (n/2)\ln(2\pi\sigma^2) - \sum_{t=1}^{n}\ln(q_t) \qquad (8)$$

where the last term is the Jacobian. Since

$$\hat{\sigma}^2 = \sum_{t=1}^{n}(\ln(q_t) - \ln(\hat{q}_t))^2 / (n-p), \qquad (9)$$

we can substitute for σ^2 in the log-likelihood to obtain the profile log-likelihood

$$\ell_p = (-n/2)\ln(\sum_{t=1}^{n}(\ln q_t - \ln \hat{q}_t)^2) - (n-p)/2 - (n/2)\ln(2\pi/(n-p)) - \sum_{t=1}^{n}\ln(q_t) \qquad (10)$$

where the last three terms are constant and can be ignored. This was maximised with a function-maximiser to estimate the β. Because the likelihood surface was bumpy, the simulated annealing version of the Nelder-Mead (1965) simplex method given by Press et al (2007) was used to maximise the likelihood function. The simplex method is already a robust method of maximisation that does not get 'stuck' on the way to the global maximum as for example a conjugate- gradient method might, and the simulated annealing modification allows it to jump over bumps. It is of course important to continue computation until one is very sure that the global maximum has been reached. To ensure this, iterations were restarted from 5 random starting points.

Another technical problem is that for some choices of the β_j, the self-consistency condition (2) cannot be satisfied. In this case, $\hat{q}$ was taken as the value of q that minimised the modulus of the difference between the left and right hand sides of (2). The solution converged such that (2) was always satisfied.

Throughout our empirical analysis, we used standard deviation rather than variance to capture risk. There is no knowledge as to which functional form is more appropriate. In general, when the functional form of the predictor is not known, it is appropriate to explore transformations of the predictor variables. We therefore fitted a model in which the standard deviation was raised to a power (effectively a Box-Cox transformation). The fit improved from that obtained with the standard deviation; the increase in log-likelihood for one extra parameter was 18, corresponding to a fall in chi-squared of 36. The power of the standard deviation was 0.766 (with standard error of 0.0123). This clearly shows that, empirically, standard deviation is a better predictor than variance. For simplicity, we have reported results using the standard deviation rather than the standard deviation raised to the power 0.766.

The calculation of standard errors on fitted model parameters, was initially done numerically by computing the Hessian of the log-likelihood using numerical differentiation, and inverting the Hessian to give the covariance matrix for the fitted model parameters. The bumpiness of the likelihood surface makes this method unreliable, so that the Hessian can have negative eigenvalues, and hence a bootstrap method was used (Efron and Tibshirani, 1993). Here the draws are resampled with replication, and the standard deviation of parameter estimates computed as the standard deviation of the distribution of resampled estimates. The 250 bootstrap samples were made by randomly selecting draws from the data set, sampling with replacement.

References

Baker R and McHale I (2009). Modelling the probability distribution of prize winnings in the UK National Lottery: consequences of conscious selection. Journal of the Royal Statistical Society, Series A 172(4): 813-834.

Brunk G G (1981). A test of the Friedman-Savage gambling model. Quarterly Journal of Economics, 96(2): 341-348.

Cook P and Clotfelter C (1993). The peculiar scale economies of lotto. American Economic Review 83(3): 634-643.

Efron B and Tibshirani R (1993). An Introduction to the Bootstrap. Chapman and Wiley: New York.

Farrell L, Morgenroth E, and Walker I (1999). A time series analysis of U.K. lottery sales: Long and short run price elasticities. Oxford Bulletin of Economics and Statistics 61(4): 513-26.

Forrest D, Gulley O D and Simmons R (2000). Testing for rational expectations in the UK National Lottery. Applied Economics 32(3): 315-326.

Friedman M and Savage L J (1948). The utility analysis of choices involving risk. Journal of Political Economy 56(4): 279-304.

Golec J and Tamarkin M (1998). Bettors love skewness, not risk, at the horse track. Journal of Political Economy 106(1): 205-225.

Greene W (2008). Econometric Analysis, 6th ed. Pearson Prentice Hall: New Jersey.

Gulley O D and Scott Jr F A (1993). The demand for wagering on state-operated lotto games. National Tax Journal 46(1): 13-22.

Kim S. and Kim H. (2016). A new metric of absolute percentage error for intermittent demand forecasts, International Journal of Forecasting, 32(3): 669-679.

Matheson V A and Grote K R (2003). Jacking up the jackpot: Are lotto consumers fooled by annuity payments? Public Finance Review 31(5): 550-567.

Murphy K and Topel R (2002). Estimation and inference in two-step econometric models. Journal of Economics and Business 20(1): 88-97.

Nelder J and Mead R (1965). A simplex method for function minimisation. Computer Journal 7(4): 308-313.

Pérez, L. and Humphreys, B. (2013). The ‘who and why’ of lottery: Empirical highlights from the seminal economic literature. Journal of Economic Surveys 27(5): 915-940.

Press W, Teukolsky S, Vetterling W and Flannery B (2007). Numerical Recipes: the Art of Scientific Computing. Cambridge University Press: Cambridge.

Sargent, T.J. (2008). ‘Rational Expectations’. In D.R. Henderson (ed.), *Concise Encyclopedia of Economics* (2nd ed.). Indianapolis: Library of Economics and Liberty. http://www.econlib.org/library/Enc/RationalExpectations.html

Scott Jr F A and Gulley O D (1995). Testing for efficiency in lotto markets. Economic Inquiry 33(2): 175-188.

Walker I and Young J (2001). An economist's guide to lottery design. Economic Journal 111(475): 700-722.