

This article was downloaded by: [University of Macau Library]

On: 30 August 2015, At: 08:37

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London, SW1P 1WG



## International Gambling Studies

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rigs20>

### What drives lottery demand? Evidence from China's lottery practice

Jia Yuan<sup>a</sup> & Jason Z. Gao<sup>b</sup>

<sup>a</sup> Faculty of Business Administration, University of Macau, Taipa, Macau

<sup>b</sup> School of Math and Statistics, Carleton University, Ottawa, Canada

Published online: 26 Feb 2015.



CrossMark

[Click for updates](#)

To cite this article: Jia Yuan & Jason Z. Gao (2015) What drives lottery demand? Evidence from China's lottery practice, *International Gambling Studies*, 15:1, 159-178, DOI: [10.1080/14459795.2014.1003316](https://doi.org/10.1080/14459795.2014.1003316)

To link to this article: <http://dx.doi.org/10.1080/14459795.2014.1003316>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms &

Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

## What drives lottery demand? Evidence from China's lottery practice

Jia Yuan<sup>a\*</sup> and Jason Z. Gao<sup>b</sup>

<sup>a</sup>*Faculty of Business Administration, University of Macau, Taipa, Macau;* <sup>b</sup>*School of Math and Statistics, Carleton University, Ottawa, Canada*

(Received 23 January 2014; accepted 22 December 2014)

As governments draw increasing revenues from the lottery industry, it has become academically important, as well as for policy purposes, to better understand the factors that can explain lottery purchase decisions. The traditional literature uses either the expected return of each lottery ticket (effective price approach) or the jackpot size (jackpot approach) to explain the variation in lottery demand. In this article, we examine these two factors by exploiting a unique lottery game set-up in lottery practice in China. This lottery game is similar to lotteries in other countries except that there is a cap policy on the grand prize, which limits the reward of each jackpot winner. We show that this complex cap policy actually causes both the lottery effective price and the jackpot size to remain almost fixed for the majority of the time while lottery demand significantly fluctuates. The lack of variation suggests that, in China's practice, neither the effective price nor the jackpot size can explain the observed variation in lottery sales. Instead, we find that the size of the lottery rollover fits well in explaining the variation in lottery demand.

**Keywords:** lottery demand; Chinese lottery

### 1. Introduction

In 2011, the global lottery market generated over \$262 billion USD in revenue and a significant portion went to various governments. For instance, in fiscal year 2011, lottery sales in the US were \$23.8 billion, and \$18.5 billion was turned over to their beneficiaries, according to LaFleur's 2012 World Lottery Almanac (LaFleur, 2012). In 2011, lottery sales in China were about \$35 billion and half became government revenue.<sup>1</sup> It has thus become important for academics, as well as for policy purposes, to understand what influences lottery purchase decisions. In this article, we attempt to understand what drives lottery demand in China's practice. We find that traditional explanations in the literature cannot explain the variation in lottery demand in China. Instead, we find that the size of the lottery rollover fits well in explaining the variation in lottery demand. To the best of our knowledge, this research is the first to establish a model and empirically examine the Chinese lottery industry in economics literature.

Traditional opinion uses the expected loss of each lottery ticket as the lottery price, which is also called the effective price, and argues that lottery demand is driven by this expected loss of betting (Cook & Clotfelter, 1993; Farrell, Morgenroth, & Walker, 1999; Forrest, Gullely, & Simmons, 2000; Rork, Fink, & Marco, 2004). Another explanation argues that lottery players dream about winning big rather than thinking about the expected return. Therefore, the jackpot size is more important for understanding lottery demand (Aruoba & Kearney, 2011; Forrest, Simmons, & Chesters, 2002).

---

\*Corresponding author. Email: [jiayuan@umac.mo](mailto:jiayuan@umac.mo)

In this article, we look at the most popular lottery game in China, which is called the Bicolor Ball Lotto (hereafter abbreviated as SSQ)<sup>2</sup> and exploit a unique policy of this lottery game. In this game, there is a cap policy on the grand prize, which limits the reward of each single winner. The grand prize winner cannot win more than a certain amount of money, as opposed to lottery games in other countries where the winner takes the entire jackpot, such as Lotto Max in Canada or Powerball in the US.

We theoretically show that this complicated cap policy actually makes the return structure surprisingly much simpler. Specifically, both the lottery price and the jackpot size remain fixed for the majority of the time while lottery demand significantly fluctuates. In other words, theoretically, neither the effective price approach nor the jackpot size approach can explain lottery demand variation in China's practice.

We then provide empirical evidence which further suggests that the lack of variation in lottery price and jackpot size cannot explain the observed variation in lottery sales in China's practice.

Instead, we provide empirical evidence which shows that lottery sales are highly correlated with the rollover money from the previous draw. We further propose and discuss the possible mechanisms through which the rollover may affect lottery demand.

This article contributes to the literature in several ways. First, to date, the literature on the modelling of demand for lotto games has focused mainly on the UK and the US. Yet, following the phenomenal growth of sales in the last decade, the Chinese lottery market has become one of the largest in the world and deserves a close examination of how it functions. To the best of our knowledge, this article is the first attempt to model the lottery demand in China. Second, this fast-growing lottery market is not merely a replicate of western lottery markets. Indeed, it has unique institutional features which provide new perspectives for us to understand the economics of lotteries. We present some results from a modelling exercise for the bestselling lotto game in China, the SSQ, and exploit some unique and interesting industry features from this Chinese lottery game. We try to gain some new insights with regard to the factors that drive lotto players to play. Third, we argue that rollover money plays a more important role in driving lottery demand than the traditional literature suggests. In the literature, rollover money affects lottery demand by entering the effective price or the jackpot size. However, in the current research, we show that rollover money may affect lottery demand over some other channels by demonstrating that the variation in either the effective price or the jackpot size is not enough to explain lottery demand fluctuation in China's practice.

Essentially, there are three potential factors which may affect lottery demand: the expected value of the lottery, jackpot size and rollover money. In western lottery practice, because there is no cap policy like China's, all three factors are entwined and people mainly use the effective price to explain lottery demand, arguably due to its simplicity and strong explanation power. However, because of the cap policy in lottery practice in China, the first two factors are not able to explain lottery demand variation and we propose that it is the third factor that explains lottery demand variation.

Nevertheless, rigorously speaking, the goal of this article is not to say that either the effective price or the jackpot size cannot influence lottery demand in China's practice if they vary. Indeed, this research presents a new world of lottery gaming where both the lottery price and jackpot size are forced to remain almost constant due to a special cap policy, and a plausible explanation on how rollover money affects lottery demand. We find that this explanation is consistent with the stylized facts in China's practice.

The rest of the article is organized as follows. Section 2 provides background information on the lottery industry and the lottery game that is examined in this study.

Section 3 is a review of the literature. Section 4 shows the calculations of the lottery price for the SSQ. Section 5 presents the data. Section 6 shows the numerical analysis and examines the relationship between rollover and lottery sales. Section 7 is a discussion of the result. Section 8 is the conclusion.

## 2. Background

### 2.1. Why China's lottery market?

The lottery market in China has become one of the largest in the world. In 2011, the aggregate lottery sales in China were over 214 billion RMB, which is around \$35 billion USD.<sup>3</sup> In this market, the SSQ is the most popular lottery game in terms of sales each week.

In fact, to the best of our knowledge, the SSQ is not only the largest lottery game in China, but also worldwide in terms of the number of tickets sold. The SSQ offers three rounds of lottery games each week. In each draw, over 150 million tickets are sold on average. The SSQ was launched in 2003 and, in the past decade, has grown dramatically. In 2011, its sales were over 48 billion RMB, which is around 7 billion USD. [Figure 1](#) shows the sales trend and the annual growth rate of the SSQ from 2003 to 2011. In those eight years, it had rapidly grown with an annual growth rate that was over 25%.

### 2.2. Background information on SSQ

The gaming rules of the SSQ are similar to those of other popular lotteries, such as Powerball in the US and Lotto Max in Canada. Each ticket is sold for 2 RMB. It requires players to pick numbers from two groups of numbers. In the first group, players need to pick 6 numbers from 1 to 33, which are called the red numbers. In the second group, players need to pick 1 number from 1 to 16, which are called the blue numbers. To win the jackpot, the player will need to match all 7 numbers randomly drawn as the winning number combination.

The SSQ has six levels of prizes, which are shown in [Table 1](#). The first prize is shown in the first row. The second prize requires the matching of all six red numbers but not the blue one. The third to sixth prizes are fixed prizes. The first and the second prizes are not fixed as the final reward depends on the number of winners and the prize pool for each payout.

[Figure 2](#) provides an even clearer illustration of the whole process of the game. When  $N$  lottery tickets are sold, the total lottery revenue is  $2N$  as each lottery ticket is sold for 2 RMB. First, 50% of the revenue from sales goes to the government and 1% goes to a so-called 'adjustment fund'. Then, the fixed prize winners take the rewards. The remaining money is called the 'high prize pool' which is reserved for the first and second prize winners. The amount of money attributed to the second prize is clearly defined as 30% of the 'high prize pool'.

The policy for the first prize is slightly more complicated. There are two scenarios. In the first case, if the rollover money from the last jackpot is less than 100 million RMB, then the grand prize winners will split the sum of the rollover money from the previous draw and the 70% from the 'high prize pool'. However, there is a cap, so if the prize is more than 5 million RMB, each winner can take only 5 million away and the rest will be rolled over to the next jackpot. The following formula illustrates this concept:

$$\text{Grand Prize from Jackpot} = \text{Min} \left\{ \frac{5\text{m}, (70\% \text{ of } H + R)}{\text{N. of Winners}} \right\}.$$

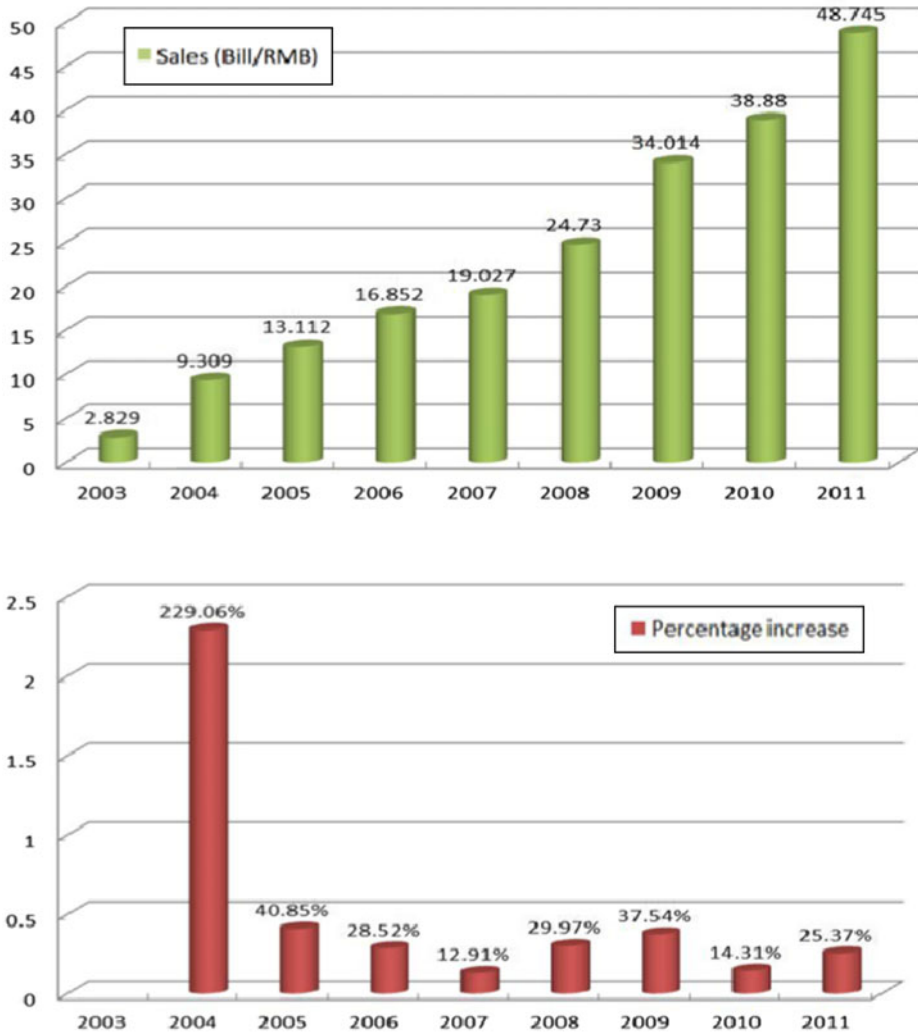


Figure 1. SSQ sales.














In the second case, if the rollover money from the last jackpot is more than 100 million RMB, the grand prize will be determined as follows:

$$\text{Grand Prize from Jackpot} = \text{Min} \left\{ \frac{5\text{m, (50\% of H + R)}}{\text{N. of Winners}} \right\} + \text{Min} \left\{ \frac{5\text{m, (50\% of H)}}{\text{N. of Winners}} \right\}$$

Winners split a two-part prize package. The first part of the jackpot money comes from the rollover from the previous jackpot and 50% of the ‘high prize pool’, and the second part is simply 20% of the ‘high prize pool’. Each of the prizes cannot exceed 5 million RMB in total.

There are two novel aspects in this lottery game. First, the jackpot has a cap, instead of using up the prize pool like other lotteries in the US and Canada. Second, this cap is

Table 1. SSQ policies.

Prize	Match		Prize distribution	Explanation
	Red balls	Blue ball		
First prize			If the rollover money from the last jackpot is less than 100 million RMB, then the grand prize jackpot winners will split the rollover from the previous draw and the 70% from the 'high prize pool'. If the prize is more than 5 million RMB, each winning ticket will only be worth 5 million RMB. If the rollover money from the last jackpot is at least 100 million RMB, there is a 2-part prize package. The winners split the rollover money from the previous draw and 50% from the 'high prize pool', as well as 20% from the 'high prize pool'. With each prize, a maximum of 5 million RMB is paid (total of 10 million RMB).	Select 6 + 1 win 6 + 1
Second prize			30% of current high prize pool	Select 6 + 1, and win 6 + 0
Third prize			Fixed amount of 3000 RMB per winning lottery ticket	Select 6 + 1, and win 5 + 1
Fourth prize	 		Fixed amount of 200 RMB per winning lottery ticket	Select 6 + 1, and win 5 + 0 or 4 + 1
Fifth prize	 		Fixed amount of 10 RMB per winning lottery ticket	Select 6 + 1, and win 4 + 0 or 3 + 1
Sixth prize			Fixed amount of 5 RMB per winning lottery ticket	Select 6 + 1, and win 2 + 1 or 1 + 1 or 0 + 1

dependent on the size of the rollover. If the rollover is less than 100 million RMB, the jackpot is no more than 5 million RMB. If the rollover money is more than 100 million RMB, the winnings are no more than 10 million RMB.

The above policy of the SSQ seems very complicated. However, in the following sections, we will show that after rigorous calculations, the results of the expected returns are surprisingly simple.

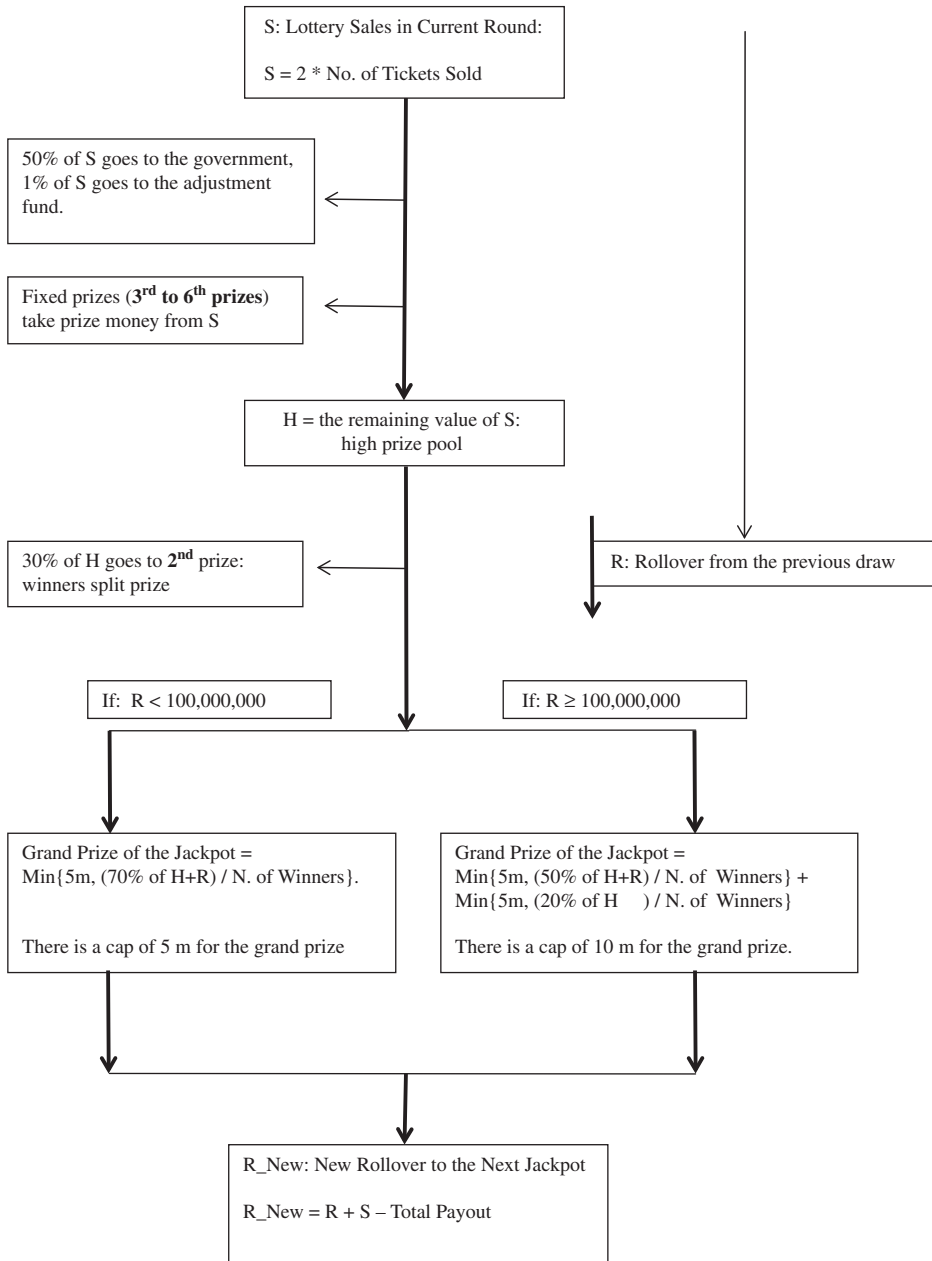


Figure 2. SSQ policies.

### 3. Related literature

Lottery games have become more and more popular worldwide and they have already become an important source of revenue for governments. There is a large volume of literature on the factors that influence lottery purchases.

One influential strand of the literature assumes that the expected loss of each lottery ticket is the major force that affects lottery demand. This expected loss for each lottery



ticket functions as the effective price of the lottery and people can rationally expect and calculate this price. Cook and Clotfelter (1993), Gulley and Scott (1993), Scoggins (1995), Scott and Gulley (1995) and Farrell and Walker (1999) are among the early researchers who established frameworks to calculate the effective price of lottery games. Thereafter, much of the research has followed their approaches to investigate various topics of lottery games. For instance, Papachristou and Karamanis (1998) calculated the effective prices of lottery games in Greece and used them to investigate the market efficiency hypothesis. Farrell et al. (1999) investigated lottery addiction by comparing the short and long run price elasticities of demand for the UK National Lottery. Forrest et al. (2000) estimated the price elasticity of lottery games in Britain and found a price elasticity close to minus one, which suggests that the British government has maximized its gaming revenue. Yuan (2011) investigated the substitution effect between two almost identical national lotteries in Canada by employing the regression discontinuity method.

However, some economists argue that lottery demand depends more on jackpot size rather than expected values because people dream of winning big. Therefore, they look at the impact of the jackpot size on lottery demand instead of the expected return of each lottery ticket. Representative works include those by Cook and Clotfelter (1993), DeBoer (1990) and Aruoba and Kearney (2011).

#### 4. Calculating effective price of SSQ

In this section, we aim to prove that under the current cap policy, the value of the expected lottery price of the SSQ is almost a constant under certain reasonable conditions, which are satisfied in the current lottery practice in China.

Let  $q_j$  denote the probability of one lottery ticket winning the  $j$ th prize. In the SSQ game, there are six levels of rewards. Therefore, we have  $1 \leq j \leq 6$ . Let  $N$  denote the total number of tickets sold. Hence,  $2N$  is the total revenue, and  $\mu_j = Nq_j$  is the average number of tickets that win the  $j$ th prize. Let  $E_j$  denote the expected return from the  $j$ th prize for each lottery ticket. For example,  $E_1$  denotes the expected return of a single lottery ticket that wins the grand prize, and  $E_2$  denotes the expected return of a single lottery ticket that wins the second prize, and so on.

Therefore, the expected return of the lottery would be  $E = E_1 + E_2 + E_3 + E_4 + E_5 + E_6$ , which will be calculated in this section.

Before we present the results on the effective price, we first summarize the expressions and values of  $q_j$  and  $E_j$  in Table 2.

It follows that the expected return from the lower or the fixed prizes for a particular player will be:

$$F = E_3 + E_4 + E_5 + E_6 \approx 0.4863.$$

According to the lottery game policy, the amount from the fixed prize winnings is subtracted from the jackpot pool and the remaining will be the prize money for the top two prizes. Therefore, the expected non-fixed prize pool, or the high prize pool, for a lottery player is  $(0.49)(2N) - FN \approx 0.494N$ .

As the lottery policy states, 30% of this high prize pool will go to the second prize. Thus, we can further calculate the expected prize pool for the second prize, which is  $0.3 \times 0.49N = 0.148N$ . Moreover, the remaining  $0.346N$  will go towards the first prize.

**Claim 1:** *The expected return of the second prize is almost fixed, which is  $E_2 \approx 0.148$  when  $N$  is over 11,000,000.*

*Proof:* See Appendix.

Table 2. Probabilities and expected returns for SSQ.

Probability	Expected return
$q_1 = \frac{1}{16 \binom{33}{6}} \approx 5.643 \times 10^{-8}$	Depends on the rollover
$q_2 = \frac{15}{16 \binom{33}{6}} \approx 8.464 \times 10^{-7}$	Depends on the rollover
$q_3 = \frac{\binom{6}{5} \binom{27}{1}}{16 \binom{33}{6}} \approx 9.142 \times 10^{-6}$	$E_3 = 3000q_3 \approx 0.0274$
$q_4 = \frac{15 \binom{6}{5} \binom{27}{1} + \binom{6}{4} \binom{27}{2}}{16 \binom{33}{6}} \approx 4.342 \times 10^{-4}$	$E_4 = 200q_4 \approx 0.0868$
$q_5 = \frac{15 \binom{6}{4} \binom{27}{2} + \binom{6}{3} \binom{27}{3}}{16 \binom{33}{6}} \approx 7.758 \times 10^{-3}$	$E_5 = 10q_5 \approx 0.0776$
$q_6 = \frac{\binom{6}{2} \binom{27}{4} + \binom{6}{1} \binom{27}{5} + \binom{27}{6}}{16 \binom{33}{6}} \approx 5.889 \times 10^{-2}$	$E_6 = 5q_6 \approx 0.2945$

Claim 1 suggests that even though the prize pool for the second prize may vary, the expected return from the second prize for any lottery player is almost fixed, which is around 15 cents.

The next claim shows that  $E_1 \approx 0.282$  when the rollover,  $R$ , is below  $10^8$ . The lower bounds on  $N$  and  $R$  are based on actual SSQ data.

*Proof:* See Appendix 1.

We also numerically calculate the values of  $E_1$  given the different possible values of the parameters. Table 3 is a record of the calculated values of  $E_1$ . It clearly shows that  $E_1 \approx 0.282$ .

One point that is worth mentioning is that  $E_1 \approx 0.282 \approx 5,000,000^*q_1$ . In other words, Claim 2 simply indicates that the jackpot prize is 5 million as a consequence of the 5 million cap policy. However, why is it that 5 million becomes the jackpot in the SSQ? The intuition is that the probability of sharing a first prize less than 5 million with others is very small, because if this happens, it means that too many people will hit the jackpot which is almost impossible to take place. For example, if the jackpot prize pool is 95 million RMB for all the jackpot winners to split, then if the final jackpot prize is less than 5 million, it means that there will be more than 20 winners to hit the jackpot at the same time. This event should be highly improbable if not impossible.

Table 3. Expected return  $E_1$  for different possible values of sales and rollover money.

		Rollover money, in 10,000,000						
		3	4	5	6	7	8	9
$N = 1.1 \times 10^8$	$E_1 =$	.2819765	.2821275	.2821473	.2821494	.2821496	.2821496	.2821496
$N = 1.2 \times 10^8$		.2819075	.2821143	.2821454	.2821492	.2821496	.2821496	.2821496
$N = 1.3 \times 10^8$		.2819053	.2821118	.2821448	.2821491	.2821495	.2821496	.2821496
$N = 1.4 \times 10^8$		.2819120	.2821110	.2821444	.2821490	.2821495	.2821496	.2821496
$N = 1.5 \times 10^8$		.2818505	.2820960	.2821415	.2821486	.2821495	.2821496	.2821496
$N = 1.6 \times 10^8$		.2818516	.2820939	.2821408	.2821484	.2821495	.2821496	.2821496
$N = 1.7 \times 10^8$		.28186241	.2820940	.2821405	.2821483	.2821495	.2821496	.2821496

Note:  $N$  is the number of players in the game. Sale =  $N * 2$  as the lottery price is 2 RMB.

The following claim deals with the case in which  $R \geq 10^8$ .

**Claim 3:** Suppose  $N > 107,000,000$  and  $R \geq 100,000,000$ . Then the expected return from the first prize is  $E_1 \approx 0.38$ .

*Proof:* See Appendix 1.

The following theorem follows immediately from Table 2 and Claims 1–3.

**Theorem 1.** Suppose  $N > 1.07 \times 10^8$  and  $R > 2.87 \times 10^7$ . Then, the effective price of

the SSQ lottery is about  $\begin{cases} 2 - (0.486 + 0.148 + 0.28) \approx 1.09 \text{ if } R < 10^8 \\ 2 - (0.486 + 0.148 + 0.38) \approx 0.99 \text{ if } R \geq 10^8 \end{cases}$ .

Theorem 1 gives the benchmark result for the analysis. The bottom line is that Chinese lottery players are facing two fixed effective prices depending on the size of rollover money from the previous round regardless how the other parameters change. This result is in contrast with the lottery practice in the Western countries without the cap policy.

Without a cap policy, the calculation of the effective price is more complicated. People need to take many parameters into consideration, such as the rollover money or the expected number of participants in the current round. Therefore, the actual effective lottery price is not known until the draw date.

In China’s practice, the calculation of the expected price is much simpler. The effective price only depends on whether the rollover size is greater than or equal to 100 million. Therefore, when lottery players observe the size of the rollover money at the beginning of each round, they do not have to do the complicated calculation that their western counterparts carry out. They will immediately know the expected price, regardless of how the other parameters change. The following section will further show that for the majority of the time, the effective price is fixed. Therefore, there is not enough variation in the effective price to explain the lottery demand variation.

### 5. Data

The data are taken directly from the SSQ authorities. The SSQ was introduced in 2003, but continued to evolve in subsequent years. In early 2009, the SSQ gaming policies were significantly modified and the version of the rules that is introduced in this article was adopted. Therefore, to avoid the influences of the structural changes of the SSQ and the learning of the SSQ players, especially the possibility that some players might not be aware of the cap policy, we choose to use data that start from 2010. The dataset contains all of the lottery sales information for 2010 and 2011. It consists of the following variables: the rollover money from the previous round, total sales in the current round, the number of

Table 4. Simple statistics for SSQ.

	Mean	S.D.	Min	Max
Sales, in <b>100,000,000</b>	2.91	0.48	2.14	4.34
Rollover in <b>100,000,000</b>	2.61	1.85	0.28	7.83
Sales change from previous draw (Absolute value)	.56%	9.97%	-35.1%	71.6%
Lottery price	1.00	0.037	0.9845	1.0837
Share of $p = 0.9845$			82.5%	
Share of $p = 1.0837$			17.5%	
Share of $p(t) = p(t - 1)$			87.8%	
Share of $p(t) \neq p(t - 1)$			12.1%	
$N = 320$				

jackpot prize winners and the single grand prize amount, the number of second prize winners and the corresponding single second prize amount, the number of winners for the other four levels of prizes, and the time. In total, we have made 320 observations.

Table 4 shows the simple statistics of the data and delivers some important information as follows. The first two rows show that the lottery sales are between 214 and 434 million RMB. As the lottery ticket price is 2 RMB, this means that the number of tickets sold is between 107 and 217 million. The second row in the table shows that the rollover money ranges between 28.7 and 783 million RMB. The last row shows that the lottery price of the SSQ is concentrated on the two values as given in our theorem.

Table 4 also shows that lottery sales are very volatile. Table 4 can be used with Figure 3 to obtain a more general picture about lottery sales. The third row of Table 4 shows the percentage of change in lottery sales compared with the last jackpot draw. In certain draws, lottery sales can drop by 35% and in others can increase by 71%. The average change in lottery sales is about .56% because positive and negative changes may cancel out each other. If we look at the average change in sales in absolute value, it is about 8%, which indicates a significant variation in lottery sales. The bottom line here is that the

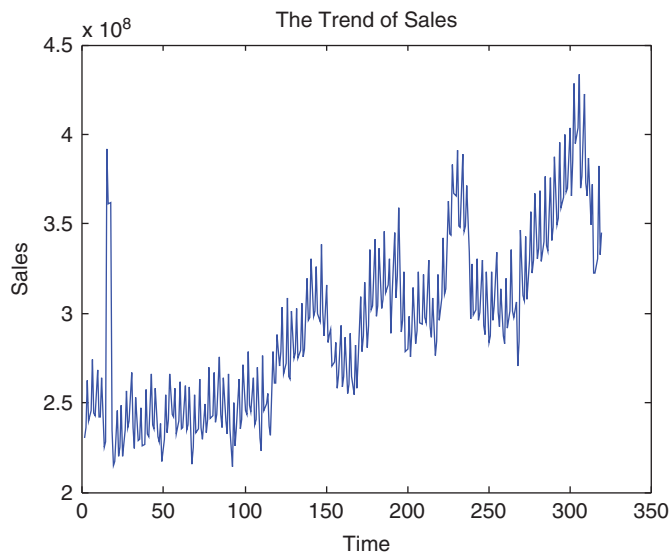


Figure 3. Lottery sales trend of the SSQ.

lottery sales of the SSQ display significant variation which warrants examination on the driving forces behind this phenomenon.

The last four rows show the calculated lottery prices for the SSQ. The results confirm the above claim. The lottery prices are very stable. The values of the lottery prices primarily are 0.9845 and 1.0837 due to the value of the rollover money. If  $R$  is less than 100 million RMB, the jackpot is capped at 5 million and if  $R$  is greater than or equal to 100 million RMB, the jackpot size will be larger and winners may possibly take home 10 million RMB. Therefore, when  $R$  is less than 100 million RMB, the lottery is more expensive as the expected return is smaller. This is 17.5% of the draws and the corresponding lottery price is around 1.0837 and the rollover size is smaller than 100 million RMB. In 82.5% of the cases, the lottery players face a more favourable return, and the lottery price is less at 0.9845 and the rollover size is greater than or equal to 100 million RMB. Overall, the average effective price in the SSQ game is 1 RMB. Compared with the nominal lottery ticket price of 2 RMB, when a lottery player decides to play this game, s/he will on average lose 50% of his/her money.

The last two rows show the portion of the time when the expected lottery prices do not change compared with the previous draw. The last two rows show that over 87% of the time, the expected lottery prices remain unchanged. This is the most direct evidence which suggests that the lottery price is not changing enough to explain the lottery sales variation.

Figure 4 further provides an examination of the relationship between lottery sales and lottery price for the SSQ. In the graph, it can be observed that lottery price is not able to explain the variation in lottery sales when it remains fixed at the concerned intervals.

We also examine the relationship between the rollover money and lottery demand. Figure 5 is a plot of the rollover money and the lottery sales in each draw. The graph shows a clear positive correlation between these two variables. A quick calculation for the correlation between sales and rollover shows that the correlation coefficient is 0.7287 by

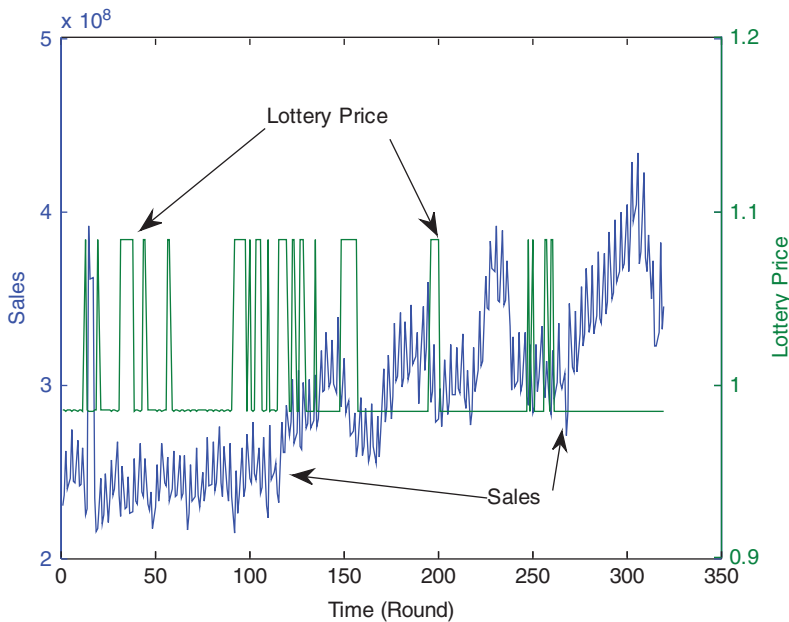


Figure 4. Lottery sales and lottery price (SSQ).

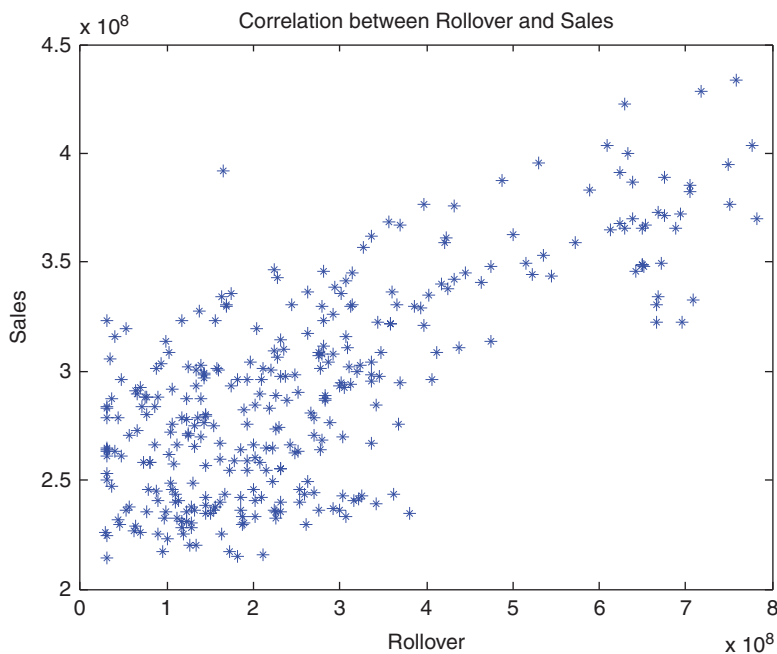


Figure 5. Rollover money and lottery sales.

using all 320 observations, and 0.7427 by removing 56 of the low rollovers. Lottery sales are highly correlated to the rollover while the effective price is almost constant.

We also examine the jackpot size distribution in the sample. In our sample, there are 56 rounds where the rollover size is smaller than 100 million RMB. The jackpot size is slightly smaller than five million RMB in only four rounds. In other words, the five-million cap policy is almost always binding, which suggests that the simple model and Claim 2 characterize real practice very well. In the sample, there are 264 rounds where the rollover size is greater than or equal to 100 million RMB. For the jackpots in these 264 rounds of the game, the mean value is 6.95 million RMB. The expected return from the average value will be  $9E_1 \approx 0.389$ ,<sup>4</sup> which is very close to the prediction of  $E_1 \approx 0.38$  as suggested by Claim 3.

To sum up, the bottom line here is that the effective price approach cannot explain the lottery demand variation in China's practice. Moreover, the rollover money is highly correlated with lottery sales, which suggests that the rollover money size is more effective in explaining the observed variation in lottery demand. In the following section, we will further investigate these arguments by running a quantitative analysis.

## 6. Regression analysis of rollover on sales

To carry out regression analysis on western lottery markets, researchers have frequently employed a two-stage least square (2SLS) approach to estimate the lottery demand function of effective price (see Forrest et al., 2000; Rork et al., 2004). A common scenario in western markets is that when lottery players purchase a lottery ticket, they do not clearly observe the effective price, given that its true value is known only when sales have closed. It is assumed that lottery players rationally expect the effective price which is also affected by lottery sales. To handle this endogeneity issue, it is necessary to use exogenous rollover size as an instrument to determine the effective price in the first stage. In fact, the use of

rollover size as an exogenous instrument has been a standard practice in estimating lottery demand in western lottery markets.

However, in current practice in China, as shown before, over 87% of the time the expected lottery prices remain unchanged. This suggests that lottery price is not changing enough to explain the lottery sales variation. Therefore, the standard empirical strategy, such as the 2SLS, is no longer considered in current practice. Instead, inspired by the relationship shown in Figure 5, we adopt the following regression to examine the relationship between the rollover and the lottery demand directly:

$$SALE = \alpha_0 + \beta Rollover + \alpha_1 DAYDUMMY + \alpha_2 Trend + \alpha_4 Sale(-i) + \varepsilon$$

Here, SALE is the lottery sales in each round. *Rollover* is the rollover money from the last draw. In western lottery markets, the day of the week is an important determinant of sales. Therefore, in the current regression, we use *DAYDUMMY* as the weekday dummy to capture the possible different impacts of the time that the lottery draw takes place. Given the phenomenal growth of lottery sales in China, we employ *Trend* to capture the time trend. We use *Sale(-i)* to represent the lagged term for the rollover money. The purpose of adopting the lagged terms of sales is to capture habit persistence as well as to help clean errors of possible autocorrelation.<sup>5</sup> We take the log value for SALE and *Rollover*.

As we are using time series data, before we run the regression, we conduct the Engle-Granger test for cointegration to rule out the possible spurious regression results. The test result rejects the spurious regression hypothesis at 99% confidence level.

Table 5 records all the regression results. In the first column, we run the regression on the subsample where there is no lottery price change. As shown in Table 4, over 87% of the time the expected lottery prices remain unchanged. However, the regression on this

Table 5. The relationship between rollover and lottery sales.

	(1) Sub-sample (no price change)	(2) Full sample	(3) Lagged terms	(4) Lagged terms	(5) Lagged terms
Rollover	0.0509*** (0.00800)	0.0519*** (0.00676)	0.0292*** (0.00668)	0.0269*** (0.00672)	0.0233*** (0.00590)
WED Dummy	0.0457*** (0.0127)	0.0493*** (0.0115)	0.0699*** (0.0107)	0.0680*** (0.0107)	0.0398*** (0.00977)
SAT Dummy	0.0416** (0.0126)	0.0432*** (0.0115)	0.0437*** (0.0104)	0.0507*** (0.0108)	0.0181 (0.00998)
TREND (in 10 <sup>-3</sup> )	1.23*** (0.0648)	1.21*** (0.0580)	0.684*** (0.0814)	0.599*** (0.0897)	0.329*** (0.0834)
Sales(-1)			0.435*** (0.0511)	0.377*** (0.0571)	0.309*** (0.0504)
Sales(-2)				0.125* (0.0553)	-0.0829 (0.0529)
Sales(-3)					0.478*** (0.0486)
CONSTANT	18.28*** (0.149)	18.26*** (0.125)	10.31*** (0.941)	9.056*** (1.085)	5.246*** (1.024)
Durbin-Watson d-statistic	.925	1.121	2.080	2.085	2.085
N	281	320	319	318	317
adj. R <sup>2</sup>	0.721	0.732	0.781	0.783	0.834

t statistics in parentheses  
\*p < 0.05, \*\*p < 0.01, \*\*\*p < 0.001

subsample shows that the rollover is highly correlated with the sales. The estimated coefficients are robust and around 0.047. In another words, the elasticity of rollover on the lottery sales is around 0.047. We further extend the regression to the full sample, and the regression results are shown in column 2. The results remain the same.

As rollover money seems to be the driving force in lottery demand variation, we further explored its regression and checked how robust the rollover can explain the lottery demand variation in China's practice. One potential issue is autocorrelation. The Durbin-Watson d-statistic in column 2 in Table 5 is around 1.121, which indicates a possible positive autocorrelation. To address this issue, in columns 3 to 5 we further add lagged terms of lottery sales to capture habit persistence as well as to help clean errors of possible autocorrelation. The results in columns 3 to 5 show that the impact of rollover on lottery sales is robust and remains the same. These estimates show that the elasticity of rollover on lottery sales is around 0.029 in the short run and 0.05 in the long run. To sum up, the impact of rollover on lottery sales is positive and significant, and the estimated coefficients are robust and around 0.05.

The regression presents some other interesting differences in comparison with western lottery markets when we examine other variables. Trend, as a variable, is often negative for many western lotteries, due to declining ticket sales for ongoing games (see DeBoer, 1990; Gulley and Scott, 1993). In China's lottery practice it plays a much larger positive role. The reason may be that the lottery market in China was experiencing rapid expansion since its legalization as opposed to the relatively stable western lottery markets. On the other hand, in western lottery markets the day of week is the single most important determinant of sales (see Forrest et al., 2000). However, in China's practice the day dummy variable does not show much difference in terms of its influence on sales. In China there are three rounds of lottery games each week and in the regression. We took Monday sales as the benchmark. There does not seem to be a significant difference between Wednesday and Saturday in terms of impact on sales.

To summarize, the regression results show that in lottery practice in China, the rollover money is highly correlated to the lottery sales data. Meanwhile, trend seems to be a more important factor that fits the data better than day of the week, as opposed to western lottery practice in which the day dummy seems to be more important.

## 7. Robustness check and discussion

One important possibility, which may potentially influence the results, is that lottery players may purchase several identical lottery tickets with the same combination of numbers in the same draw, because they hope to win the '5 million' or '10 million' jackpot more than once in a single draw. For instance, some lottery players may purchase two lottery tickets at the same time with the same combination of lottery numbers. Then, if those numbers hit the jackpot, s/he can win two grand prizes.

If a significant number of players adopt this strategy, then the function that we use to numerically calculate the lottery price for  $E_1$  will change dramatically. Then it is possible for effective lottery prices to significantly vary. Moreover, if the distribution of the number of repeated lottery purchases varies from draw to draw, this may also cause the lottery price to vary.

Therefore, we need to examine the distribution of the lottery purchase. What is the portion of lottery tickets purchased in which one ticket is bought for one set of numbers? What is the portion of lottery tickets purchased in which two tickets are bought for the same combination of numbers? And so on and so forth.



We have obtained a unique lottery dataset that contains the information on how individual lottery players choose their lottery numbers for the SSQ. The data directly originate from the Taobao Lottery, which is the largest online platform for lottery purchases in China. Specifically, the Taobao Lottery provides an online platform which lottery players can use to purchase the SSQ. Of course, it is impossible for us to observe all the lottery number selections in the SSQ game. However, we argue that this dataset is good enough to examine how lottery players pick numbers and purchase lottery tickets in this research.

The dataset contains 30,366 SSQ lottery players and 3,557,606 purchased tickets. It is a record of all the lottery number selection information from draws #2011141 to #2011149 in 2011. We can observe whether an individual lottery player purchased two or three lottery tickets for the same combination of numbers.

Table 6 illustrates the distribution of lottery purchase strategies and the distribution variation for each draw. There are two important issues that are worth discussing here. First, it is shown that around 98% of lottery players purchase one ticket for the same combination of lottery numbers. Around 1% of players purchase two tickets for the same combination of numbers. This means that a lottery player does not have to worry about sharing the jackpot with a repetitive lottery player who purchases many tickets with the same numbers. In other words, it is reasonable to use the function in Section 3 to calculate the lottery prizes.

Second, the distribution does not significantly vary. For example, in draw 2011141, a 'single ticket' constitutes about 98.2% of the entire sample and does not change much across all nine draws. The share of 'double purchases', where an individual player purchases the same combination of numbers twice, is also very stable, at around 1.2%. To summarize, the distribution of the lottery purchase strategy is stable. Of course, it is possible that a lottery player may increase repeated purchases for the same combination of numbers due to increases in the prize funds, but the entry of new lottery players who only purchase one single ticket may still maintain a stable lottery strategy distribution, as shown in Table 6.

In other words, when an individual lottery player purchases one ticket, the expected return of this single ticket can be calculated by using the function in Section 3, which shows that this return is almost constant.

Table 6. Distribution of lottery purchase strategies.

Round =	Shares (in %)			N of tickets	N of players
	Single	Double	More than 3		
2011141	98.2%	1.23%	0.56%	360,432	12,968
2011142	98.0%	1.45%	0.55%	368,265	13,462
2011143	98.4%	1.04%	0.59%	371,667	12,653
2011144	98.2%	1.2%	0.65%	393,255	12,620
2011145	98.1%	1.2%	0.64%	395,660	12,866
2011146	98.0%	1.2%	0.71%	397,808	12,150
2011147	98.2%	1.2%	0.65%	398,376	12,450
2011148	98.2%	1.2%	0.61%	423,067	12,857
2011149	98.5%	0.9%	0.56%	449,076	12,074
<b>Total number of tickets</b>				3,557,606	
<b>Total number of players</b>				30,366	

Data source: Taobao Lottery.

Another issue is that if rollovers are an important driving force for lottery demand in China's practice, what are the possible mechanisms behind this phenomenon? There may be many possible mechanisms. First, higher rollovers may cause greater advertising impacts in terms of the media. People may be more easily influenced through media broadcasts on the higher rollover money. Unfortunately, there is no available related data in the lottery industry and it is difficult to test this conjecture.

It is also possible that rollover money size may affect lottery demand by taking into consideration the subjective probability perceived by lottery players in winning the jackpot. In the standard expected utility model, the weight placed on the jackpot is actually the probability of winning the jackpot. In the SSQ game, this probability weight is  $5.643 \times 10^{-8}$ . However, we argue that few people have a real sense of what this probability actually means. Instead of using this probability, lottery players may assign a subjective probability weight for the jackpot. It is possible that the rollover money drives the lottery demand by influencing lottery players with higher amounts of rollover money, thus causing more weight to be placed on the jackpot through subjective probability.

Then, one would further ask: what is the mechanism that would cause higher rollover money to increase the subjective probability weight assigned by a lottery player? One intuition would be as follows. When a lottery player observes that the rollover money flows into the current lottery round, s/he may form the belief that 'I am the one chosen by God to claim the jackpot.' When the rollover money grows larger in size, this will further reinforce his/her belief or, in other words, increase his/her subjectiveness of the probability that s/he will be 'the one' to win the jackpot.

To summarize, there are essentially three potential factors that may affect lottery demand: the expected value of the lottery, jackpot size and rollover money. In western lottery practice, because there is no cap policy, which is unlike the situation in China, all three factors are entwined and people mainly use the effective price to explain lottery demand, arguably due to its simplicity and strong explanation power. However, because of the cap policy in lottery practice in China, the first two factors are not able to explain the lottery demand variation and we propose this third factor as an explanation.

## 8. Conclusion

As governments draw increasing revenues from the lottery industry, it has become important to understand the factors that can explain lottery purchase decisions. However, to explain lottery betting behaviour is always a challenging task. Most of the traditional research has mainly focused on western lottery practices, and the traditional literature mainly uses either the expected return of each lottery ticket (effective price approach) or the jackpot size (jackpot approach) to explain the variation in lottery demand.

The lottery practice in China provides researchers with a new perspective in understanding the economics of lotteries because in China the lottery authorities have put many different but interesting policies into practice, such as the cap policy. By studying the most popular lotto game in China and exploiting the unique cap policy, we show that this complex cap policy actually causes both the lottery effective price and the jackpot size to remain almost fixed for the majority of the time while lottery demand significantly fluctuates. The lack of variation suggests that in China's practice, neither the effective price nor the jackpot size can explain the observed variation in lottery sales. Instead, we find that the size of the lottery rollover fits well in explaining the variation in lottery demand.

Essentially, all the three factors – the expected value of the lottery, jackpot size and rollover money – may potentially affect lottery demand. However, in western lottery

practice, all three factors are entwined and fluctuate together. Thanks to the unique cap policy in China, by keeping the other two factors fixed, we present evidence to show that lottery rollover alone can affect lottery demand.

If rollovers are an important driving force for lottery demand, what are the possible mechanisms behind it? Due to the limitation of the data, in the current study we cannot directly test all the possible conjectures such as advertising factors, culture difference factors or behavioural factors. This is beyond scope of the current study and will be left for future work.

### Acknowledgements

We are indebted to the editor of this journal and the three anonymous referees for helping us to substantially improve the paper. We would also like to thank Travis Ng, Li Tao, Sun Guangzhen and the seminar participants at the Fudan University, University of Hong Kong and the First Conference of the APCG (Asia Pacific Association for Gambling Studies), 2012 for comments and suggestions. All errors are our own.

### Funding

This work was supported by the University of Macau [grant number MYRG006-FBA13].

### Conflicts of interest

Funding sources: None.

Competing interests: No potential conflicts of interest were reported by the authors. Both authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

### Constraints on publishing

None.

### Notes

1. In 2011, the total government revenue in China was around \$1.5 trillion RMB (*China Annual Statistic Yearbook*, 2011). The revenue from the lottery games in China was around three times that from the securities transaction stamp tax in 2011.
2. The nominal price for each lottery ticket is 2 RMB. In each draw, over 150 million lottery tickets are sold on average and the average lottery revenue is over 300 million RMB. The rules of the game are very similar to other renowned lottery games worldwide, such as Lotto Max in Canada or Powerball in the US. The jackpot requires the matching of 7 numbers out of 43 which are randomly drawn. The abbreviation used for the Bicolor Ball Lotto is SSQ, following the initials of its Chinese phonetic spelling. More details of the SSQ will be discussed in Section 2.
3. The size of China's lottery market is comparable with that of other major lottery markets. For instance, in fiscal year 2011, lottery sales in the US were \$23.8 billion according to LaFleur's 2012 World Lottery Almanac (LaFleur, 2012).
4.  $9E_1 \approx 0.389 \approx 6,950,000^*q_1$ .
5. We thank one of the referees for this suggestion.

### Notes on contributors

Jia Yuan received his PhD in Economics in 2009 from the University of Minnesota. He is currently an assistant professor at the Faculty of Business Administration at the University of Macau. His research interests include empirical industrial organization, gambling study, and applied economics.

He has published in the *Journal of Gambling Studies*, *Journal of Gambling Business and Economics*, etc.

Jason Z. Gao received his PhD in 1989 from University of California at San Diego. He is currently a full professor at the School of Mathematics and Statistics, Carleton University, Ottawa, Canada. His research focuses on discrete mathematics, mathematics of gambling, and game theory. He also does consulting work for the gambling industry and telecommunication industry.

## References

- Aruoba, B., & Kearney, M. (2011). Do lottery gamblers love risk or overweight small odds? Working Paper.
- China Annual Statistic Yearbook, 2011. Beijing: China Statistics Press.
- Cook, P., & Clotfelter, C. (1993, June). The peculiar scale economies of lotto. *American Economic Review*, 83, 634–643.
- DeBoer, L. (1990). Lotto sales stagnation: Product maturity or small jackpots? *Growth and Change*, 21, 73–77.
- Farrell, L., Morgenroth, E., & Walker, I. (1999). A time series analysis of UK lottery sales: Long and short run price elasticities. *Oxford Bulletin of Economics and Statistics*, 61, 513–526.
- Farrell, L., & Walker, I. (1999). The welfare effects of lotto: Evidence from the UK. *Journal of Public Economic*, 72, 99–120.
- Forrest, D., Gulley, D., & Simmons, R. (2000). Elasticity of demand for UK national lottery tickets. *National Tax Journal*, 4, 853–863.
- Forrest, D., Simmons, R., & Chesters, N. (2002). Buying a dream: Alternative models of demand for lotto. *Economic Inquiry*, 40, 485–496.
- Gulley, D., & Scott, F. (1993). The demand for wagering on state operated lotto game. *National Tax Journal*, 46, 13–22.
- LaFleur, T. (2012). LaFleur's 2012 World Lottery Almanac and lottery annual reports. TLF Publications, Inc.
- Papachristou, G., & Karamanis, D. (1998). Investigating efficiency in betting markets: Evidence from the Greek 6/49 Lotto. *Journal of Banking and Finance*, 22, 1597–1615.
- Rork, J., Fink, S., & Marco, A. (2004). Lotto nothing? The budgetary response to state lotteries. *Applied Economics*, 36, 2357–2367.
- Scoggins, J. (1995). The lotto and expected net revenue. *National Tax Journal*, 48, 61–70.
- Scott, F., & Gulley, D. (1995, April). Testing efficiency in lotto markets. *Economic Inquiry*, 33, 175–188.
- Yuan, J. (2011). Are lotteries substitutes for each other? *Journal of Gambling Business and Economics*, 5(3), 1–14.

## Appendix 1

### Proof for Claim 1

**Proof.** We assume that the players independently choose their numbers. It is easy to show that the probability of a player and  $k$  other players who will win the second prize is

$$q_2 \binom{N-1}{k} q_2^k (1-q_2)^{N-1-k}.$$

Hence, the expected return of a lottery player with the 2nd prize is

$$E_2 \approx \sum_{k=0}^{N-1} \frac{0.148N}{k+1} \binom{N-1}{k} q_2^k (1-q_2)^{N-1-k} q_2 = \frac{0.148N}{N} (1 - (1-q_2)^N) \approx 0.148. \quad (1)$$

Here, we use the following two facts:

$$\sum_{k=0}^{N-1} \frac{1}{k+1} \binom{N-1}{k} x^{k+1} y^{N-1-k} = \frac{1}{N} (x + y^N - y^N) \tag{2}$$

and  $(1 - q_2)^N < 0.001$  when  $N \geq 11,000,000$ . **QED**

**Proof for Claim 2**

**Proof.** We first note that the first prize pool is  $R + (0.70)(0.494N) = R + 0.346N$ , which is the summation of the rollover money and the money left after taking away the fixed prize amounts and the second prize pool. We also note that, in this case, the prize is capped at 5 million RMB.

We aim to calculate the expected return of one single ticket that wins the grand prize. Suppose this winning player and  $k$  other players share the grand prize. Then, his/her share of the grand prize will be  $\min\left\{5000000, \frac{R+0.346N}{k+1}\right\}$ . By following the same argument per the proof for Claim 1, we can show that his/her expected return from the grand prize is

$$E_1 = \sum_{k=0}^{N-1} \min\left\{5,000,000, \frac{R + 0.346N}{k + 1}\right\} \binom{N-1}{k} q_1^k (1 - q_1)^{N-1-k} q_1 \tag{3}$$

The above equation shows that  $E_1$  is a function of the rollover money,  $R$ , and the number of tickets sold,  $N$ .

By assuming  $N > 1.07 \times 10^8$  and  $R > 2.87 \times 10^7$ , we have  $\frac{R+0.346N}{13} > 5 \times 10^6$ , and hence,  $E_1 > \sum_{k=0}^{12} 5 \times 10^6 \binom{N-1}{k} q_1^k (1 - q_1)^{N-1-k} q_1 > 5 \times 10^6 q_1 \times 0.99 = 0.280$ .

It is clear from (3) that  $E_1 \leq 5 \times 10^6 q_1 \approx 0.282$ . Hence  $0.280 \leq E_1 \leq 0.282$ . **QED**

More accurate values of  $E_1$  can be obtained by using (3) and Poisson approximation:

$$\binom{N-1}{k} q_1^k (1 - q_1)^{N-1-k} \approx e^{-\mu} \frac{\mu^k}{k!}, \quad \text{where } \mu = (N - 1)q_1. \tag{4}$$

Let  $n$  be the greatest integer such that  $\frac{R+0.346N}{n} \geq 5 \times 10^6$ . Then it follows from (3) and (4) that

$$\begin{aligned} E_1 &= \sum_{k=0}^{n-1} 5 \times 10^6 q_1 e^{-\mu} \frac{\mu^k}{k!} + \sum_{k=n}^{N-1} \frac{R + 0.346N}{k + 1} q_1 e^{-\mu} \frac{\mu^k}{k!} \\ &\approx 0.282 e^{-\mu} \sum_{k=0}^{n-1} \frac{\mu^k}{k!} + \frac{R + 0.346N}{\mu} q_1 \left(1 - \sum_{k=0}^n e^{-\mu} \frac{\mu^k}{k!}\right) \end{aligned} \tag{5}$$

For example, when  $N = 1.07 \times 10^8$  and  $R = 2.87 \times 10^7$ , we have  $\mu \approx 6.038$  and  $n = 13$ , and hence  $E_1 \approx 0.280 + 0.002 = 0.282$

**Proof for Claim 3**

**Proof.** The proof is similar to that of Claim 2. Here, we have

$$\begin{aligned} E_1 &= \sum_{k=0}^{N-1} \min\left\{5 \times 10^6, \frac{R + 0.247N}{k + 1}\right\} \binom{N-1}{k} q_1^k (1 - q_1)^{N-1-k} q_1 \\ &\quad + \sum_{k=0}^{N-1} \min\left\{5 \times 10^6, \frac{0.099N}{k + 1}\right\} \binom{N-1}{k} q_1^k (1 - q_1)^{N-1-k} q_1 \end{aligned}$$

Downloaded by [University of Macau Library] at 08:37 30 August 2015

When  $N > 1.07 \times 10^8$  and  $R \geq 10^8$ , we have

$$\begin{aligned} \frac{R + 0.247N}{25} &> 5 \times 10^6 \text{ and } \frac{R + 0.247N}{26} < 5 \times 10^6 \\ \sum_{k=0}^{N-1} \min \left\{ 5 \times 10^6, \frac{R + 0.247N}{k+1} \right\} \binom{N-1}{k} q_1^k (1-q_1)^{N-1-k} q_1 \\ &\approx \sum_{k=0}^{24} 5 \times 10^6 \binom{N-1}{k} q_1^k (1-q_1)^{N-1-k} q_1 \approx 0.282. \end{aligned}$$

Let  $n$  be the largest integer such that  $\frac{0.099N}{n} > 5 \times 10^6$ . Then

$$\begin{aligned} E_1 &\approx 0.282 + \sum_{k=0}^{N-1} \min \left\{ 5 \times 10^6, \frac{R + 0.247N}{k+1} \right\} \binom{N-1}{k} q_1^k (1-q_1)^{N-1-k} q_1 \\ &= 0.282 + \sum_{k=0}^{n-1} 5 \times 10^6 \binom{N-1}{k} q_1^k (1-q_1)^{N-1-k} q_1 \\ &\quad + \sum_{k=n}^{N-1} \frac{0.099N}{k+1} \binom{N-1}{k} q_1^k (1-q_1)^{N-1-k} q_1 \\ &\approx 0.282 + 0.0282 \sum_{k=0}^{n-1} e^{-\mu} \frac{\mu^k}{k!} + 0.099 \left( 1 - \sum_{k=0}^n e^{-\mu} \frac{\mu^k}{k!} \right) \\ &\approx 0.381 + 0.183 \sum_{k=0}^{n-1} e^{-\mu} \frac{\mu^k}{k!} - 0.099 e^{-\mu} \frac{\mu^n}{n!} \end{aligned}$$

For example, when  $N = 1.07 \times 10^8$  (a lower bound of N under consideration), we have  $\mu \approx 6.038$  and  $n = 2$ , and hence

$$E_1 \approx 0.381 + 0.183 e^{-\mu} (1 + \mu) - 0.099 e^{-\mu} \frac{\mu^2}{2} \approx 0.380$$

When  $N = 2.17 \times 10^8$  (an upper bound of N under consideration), we have  $\mu \approx 12.245$  and  $n = 4$ , and hence

$$E_1 \approx 0.381 + 0.183 e^{-\mu} \left( 1 + \mu + \frac{\mu^2}{2} + \frac{\mu^3}{6} \right) - 0.099 e^{-\mu} \frac{\mu^4}{24} \approx 0.381$$

QED