Lottery Markets–Design, Micro-Structure, and Macro-Behavior: An Agent-Based Computational Approach

Shu-Heng Chen
AI-ECON Research Center
Department of Economics
National Chengchi University
Taipei, Taiwan 11623
chchen@nccu.edu.tw

Bin-Tzong Chie
AI-ECON Research Center
Department of Economics
National Chengchi University
Taipei, Taiwan 11623
chie@aiecon.org

Chia-Wei Lee
AI-ECON Research Center
Department of Economics
National Chengchi University
Taipei, Taiwan 11623
dae0504@yahoo.com.tw

Abstract- Given the failure to generate the prevalence phenomenon known as the halo effect in the lottery market, this paper proposes a new design of the agent-based computational lottery market based on the expected-utility maximization paradigm. In this paper, we shall demonstrate the performance of the lottery market with this new design.

1 Motivation and Literature Review

Chen and Chie (2003) pioneer an agent-based computational model of the lottery market. Using the model, one can simulate and analyze different consequences of different designs of lottery markets. For example, the specific question addressed in that paper is the *optimal lottery tax rate*, and, as one of their main experiment demonstrates, the optimal tax rate is found to be around 40%, which interestingly mimics empirical observations. Despite this interesting finding, their analysis does encounter one major weakness, i.e., the absence of the *halo effect*.[1]

One possible explanation to the absence of the halo effect is that agents' behavior in the original model may be a little overwhelming in the sense that it leave agents too many parameters to learn. With this degree of sophistication, it becomes difficult for agents to make sense of the simple relation between the *jackpot size* and *lottery participation*, which may indeed do not have any simple rationale behind it. Therefore, in this paper, we try to model agent in a simpler way, which is very much motivated by the standard *expected-utility maximization* paradigm. We are well aware that the expected-utility maximization will not entice agents to purchase lottery tickets, because the expected return is normally negative for lottery investment. Nevertheless, if agents have expectations *differing* from the objective odds of winning, then it is still possible to have lottery participation. Furthermore, given agents' subjective beliefs, when the jackpot size increases, agents will generally intensify their lottery participation. As a result, the halo effect or lottomania is embedded in agents' behavior.[2]

As said, this new approach heavily depends on agents' subjective beliefs, which shall be, at least initially, heterogeneous among agents. Nonetheless, as agents are adaptive or learning, these subjective beliefs will evolve over time. Then it leave us the issue whether the phenomenon of lottomania can survive in this situation.[3]

In sum, despite the use of the expected-utility maximization framework, the analysis is still in spirit of *bounded rationality* in the sense that agents' beliefs are adaptive, specifically, they are adapted with *genetic algorithms*.

[1] The phenomenon that sales following a rollover are higher than sales prior to the rollover is known in the industry as the *halo effect* ([2], [3]). The halo effect is partially due to considerate media attention paid to rollovers, which in turn creates a bout of *lottomania*.

[2] Of course, the disadvantage is also clear; the lottomania, if it happens, can no longer be considered as an emerging property.

2 Agent-Based Modeling of the Lottery Market

We shall follow Chen and Chie (2003) to build up the three main aspects of lottery behavior. The three main aspects are *lottomania*, *conscious selection*, and *aversion to regret*.

The difference between this version of agent-based lottery market and the early one in [1] mainly lies in the formation of the decision on *lottery participation*, α, which is defined as the percentage of the income spent on lottery purchasing. For the latter, it is formulated as a standard fuzzy decision rule on which agents base their expenditures of purchasing lottery tickets, and the key variable in the fuzzy decision rule is the size of the jackpot. Agents' utility, however, is not directly involved. In this version, the decision of the lottery participation is made explicitly within an expected-utility framework.

Let p_i be agent i's *subjective belief* (probability) that he will win the jackpot. We index p_i by t as $p_{i,t}$ when we want to emphasize its adaptation and dynamics. With the device of the expected-utility maximization, agents' decision on lottery participation α_i can be treated as a control variable to solve the following optimization problem.

$$\max_{\alpha_i \in \{0,1\}} u_i = \max_{\alpha_i \in \{0,1\}} (1-p_i)u_i[(1-\alpha_i)I] + p_i u_i[((1-\alpha_i)I + J], \tag{1}$$

where J denotes the jackpot size and I denotes the income agent received in each period.[4] Chen and Chie (2003) assume all agents are risk-neutral, and $u(c) = c$. In this paper,

[3] As has been well argued in [1], lottomania is not an exogenously setting, it is an endogenously generated because one has no control over the jackpot size, which in turn is an aggregate outcome from individuals' decision on lottery participation.

[4] We notice that this is not a precise formulation of the expected utility since there are several different prizes associated with different winning probabilities. For simplicity, only the jackpot prize is considered in this paper. In this paper, income I is exogenously given, and is identical for all agents. In a separate paper, Chen, Chi and Fan (2005) study the case that income is heterogeneous among agents, and examine the effect of income distribution on lottery participation.

we follow a more standard formulation in finance, i.e., to assume that agents are *risk-averse* and hence $u(c) = \log c$.

It is then easy to show the optimal solution for the problem (1) is

$$\alpha_i^* = \begin{cases} \frac{I-(\frac{I}{J^{p_i}})^{\frac{1}{1-p_i}}}{I}, & if \quad \log^I \leq p_i \log^J + (1-p_i)\log^{(I-e)}, \\ 0, & otherwise. \end{cases} \tag{2}$$

where e is unit cost of the lottery ticket.

The final *fitness function* after taking *aversion to regret*, θ_i, into account then is

$$u_i(c) = \begin{cases} u_i[(1-\theta_i)c], & if \ \ \alpha_i^* = 0 \ \ and \ \ N_x > 0, \\ u_i[(1+\theta_i)c], & if \ \ \alpha_i^* = 0 \ \ and \ \ N_x = 0, \\ u_i[(1-\alpha_i^*)I + \pi_i], & \quad otherwise. \end{cases} \tag{3}$$

The θ_i in the utility function (3) measures how regretful the non-participant, characterized by $\alpha_i^* = 0$, would be if the jackpot is drawn i.e., the number of winners of the jackpot is positive ($N_x > 0$). On the other hand, opposite to regret, the non-gamblers may also derive pleasure from gamblers' misfortune, in particular when the jackpot is not drawn ($N_x = 0$). Since generally mass media only gives a large converge on the jackpot winners, not the losers, we therefore assume that the regret effect is asymmetric between lottery non-participants and participants, as shown in the last part of Equation (3).

The last feature of our model of agents is the *conscious selection*, which refers to non-random selections of the combination of numbers. To take conscious selection into account, let $\vec{b}_i$ be a X-dimensional vector, whose entities take either "0" or "1". Consider a number z, where $1 \leq z \leq X$. If "0" appears in the respective zth dimension, that means the number z will not be consciously selected by the agent, while "1" indicates the opposite. If $\vec{b}_i$ has exactly x 1s, then one and only one combination is defined and the agent would select only that combination while purchasing the lottery ticket(s). If $\vec{b}_i$ has more than x 1s, then many more combinations can be defined. The agent will then randomly select from these combinations, while purchasing the ticket(s). Finally, if $\vec{b}_i$ has less than x 1s, then those designated numbers will appear in each ticket bought by the agent, whereas the rest will be randomly selected from the non-designated numbers.

All these three aspects of lottery behavior will adapt and change over time; therefore, they are all index by t as $p_{i,t}, \theta_{i,t}$, and $\vec{b}_{i,t}$. The evolution of the lottery behavior is then driven genetic algorithms.

3 Genetic Algorithms

The standard genetic algorithm is applied to evolving the three behavioral parameters of the entire population, $POP_t \equiv \{p_{i,t}, \vec{b}_{i,t}, \theta_{i,t}\}_{i=1}^N$. GA starts with encoding the behavior parameters to binary digits (binary coding) or real numbers (real coding), usually called the *chromosome* (*finite-string*) representation. Here, we apply *real coding* to $p_{i,t}$ and $\theta_{i,t}$, since they are real numbers between 0 and 1. However, binary coding is applied to $\vec{b}_{i,t}$ given that it is a binary vector.

Table 1: Experimental Design

Market Parameters	
Pick x from X (x/X)	5/16
Lottery Tax Rate (τ)	10%, ..., 90%
$s_0, s_1, ..., s_5$	0%, 0%, 35%, 15%, 12%, 38%
Drawing Periods ($\bar{r}$)	3
Number of Agents (N)	5000
Income (I)	200
GA Parameters	
Range of $p_{i,0}$	[0, 0.003]
Periods (Generations) (T)	500
Crossover Rate (P_c)	90%
Mutation Rate (P_m)	0.1%
Arithmetic Mutation Size	Equation (4)
Tournament Size (φ)	200
Generation Gap (η)	100

Tournament selection, associated with a *tournament size* φ, is employed as the selection scheme to determine who are the celebrities (mating pool). Given the mating pool, offspring are generated by applying the two genetic operators: crossover and mutation. First, the *crossover*. Since each chromosome represents the three different aspects of agents' behavior, the crossover is made in a pair-by-pair manner, i.e., to restrict the swap only to the paired characteristic, called *paired crossover*. Each time crossover work on only one of the three pairs which are determined randomly. If it is the pair of $\vec{b}_{i,t}$, the one-point crossover with a crossover rate P_c is applied. If it is $p_{i,t}$ or $\theta_{i,t}$ the arithmetic crossover with the same crossover rate is applied. Second, the mutation. After crossover, each part of the resultant chromosome has a chance of being mutated. For $\vec{b}_{i,t}$, the bit mutation with a mutation rate P_m is applied, whereas for $p_{i,t}$ or $\theta_{i,t}$, an arithmetic mutation adjusted to an equivalent bit-mutation is applied to these real parameters as Equation (4).[5]

$$p^{new} = p^{old} + \sum_{i=1}^{16} B_{P_m} (\frac{1}{2})^i \cdot (-1)^{B_{\frac{1}{2}}}, \tag{4}$$

where p_i^{old} and p_i^{new} indicates the subject belief *before* and *after*. B_{P_m} and $B_{\frac{1}{2}}$ are the Bernoulli random variables with success probability P_m, i.e, the mutation rate, and one half respectively.

When all offspring are produced, the steady-state replacement with the generation gap η is applied to replace the old generation. By the parameter η, the agents belonging to the top $1 - \eta$ per cent would remains, and only the agents belonging to the bottom η percent would be replaced by offspring.

4 Experiment Designs

We would like to see how this new design may come up with anything significantly different from what was obtained in [1]. The behavior of the lottery market studied in [1] includes the optimal lottery tax rate associated with a simu-

[5] As the arithmetic crossover, we could use the standard arithmetic mutation for the real-coding chromosomes. The reason using this bit-mutation design is to make our results comparable to [1].

lated Laffer curve, the impact of the regret effect upon the Laffer curve, the statistical relation between *rollovers* and *sales* (the halo effect), the evolution of conscious-selection behavior and the inter-dependence preference (the aversion to regret).

To be able to compare the results of our new design with those of Chen and Chie (2003), we follow almost the same experimental design as theirs (see Table 1). The design is composing of two parts, namely, *market parameters* and *GA parameters*.

It is anticipated that s_y is increasing in y. To recap, a lottery game can be represented by the following x+4-tuple vector: $\mathcal{L} = (x, X, \tau, s_0, ..., s_x)$, which is also shown in the upper half of Table 1.

5 Simulation Results

Twenty five independent runs are conducted with the design as indicated in Table 1. The results presented below is therefore not based on a single run, but statistics of these 25 runs.

5.1 Lottery Tax Rate and Tax Revenue

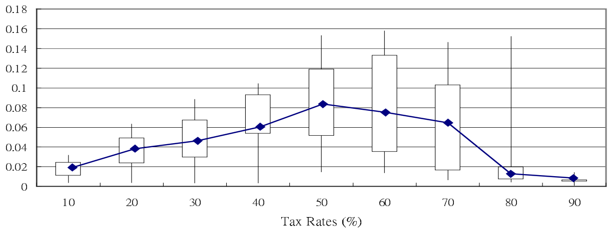

Figure 1: Tax Revenue Curve and the Associated Box-Whisker Plot

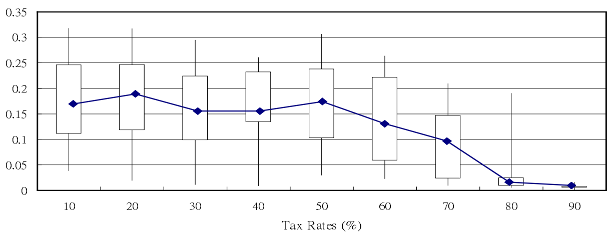

Figure 2: Lottery Participation Rate (α)

Figure 1 is the *Laffer curve* phenomenon noticed by Chen and Chie (2003). The result here is comparable to the one observed earlier. The most noticeable difference is the location of peak. For Chen and Chie (2003), it is at a tax rate of 40%, but this one is higher up to 50%. In addition to shift of the peak, we also see the change in distribution of the effective tax rate. By looking at the range (the box-and-whisker plot) of the effective tax rate, there is generally an observed downward tendency. [6] Figure 2, however, interestingly shows that the lottery participation rate (α) is actually not very sensitive to the lottery tax rate within a certain rage up to 50%.

5.2 Rollovers and Sales

Table 2: Statistics of Rollover and Sales

Tax Rates	t statistic (p-value)	α_1 (p-value)	R^2	anomalies
0.1	-92.73 (0.0000)	0.35 (0.0000)	0.1065	25%
0.2	-89.36 (0.0000)	0.35 (0.0000)	0.1070	22%
0.3	-90.40 (0.0000)	0.22 (0.0000)	0.0813	25%
0.4	-88.97 (0.0000)	0.29 (0.0000)	0.0987	23%
0.5	-98.10 (0.0000)	0.15 (0.0000)	0.0558	22%
0.6	-86.30 (0.0000)	0.16 (0.0000)	0.0715	27%
0.7	-61.58 (0.0000)	0.16 (0.0000)	0.0705	34%
0.8	-36.09 (0.0000)	0.07 (0.0000)	0.0366	42%
0.9	-31.86 (0.0000)	0.02 (0.0000)	0.0402	42%

It is generally assumed that large size of rollovers will enhance the attractiveness of the lottery game. Statistics also tell us that the mean sales conditional on the rollover draw is normally *higher* than that of the regular draw. Nevertheless, this phenomenon is not able to be fully replicated in [1]. This becomes a puzzle, and they call it *disappearance of the halo effect*. A conjecture of this failure has been given in the introduction section of the paper, which also motivates a different design of agent's behavior in this paper. But, would this new design be able to deliver the halo effect?

To answer the question, we perform similar statistical tests as what Chen and Chie (2003) have done, and these are shown in Table 2. The t statistic shown in the second column is a test statistic for the null that the mean sales of the rollover draw is greater than that the regular draw, i.e, the halo effect. From the corresponding p value, we can see that the halo effect is again uniformly rejected. This result is consistent with what was found in [1]: *the halo effect is again absent*.

But, sales may actually fall in some rollover draws, and the frequency of this *anomaly* can be high up to 20% to 25% in some countries. However, the frequency of anomalies found in [1] is around 60% (τ=0.4, 0.5), which is simply too high to be comparable with the real data. Nevertheless, the fifth column of Table 2 does pin down out a more realistic frequency of anomalies, say 23% (τ=0.4) or 22%(τ=0.5). In fact, this is exactly the same number we observed from the U.K data.Therefore, while this new design does not completely cure the problem, it does make some improvement in terms of replicating the real-world data. [7]

[6] The appearance of Laffer curve basically shows the two counterbalancing forces as shown in Equations (5) and (6).

$$T = tS = t(\alpha I) \tag{5}$$

$$\frac{\partial T}{\partial t} = \underbrace{S}_{+} + t \underbrace{\frac{\partial \alpha}{\partial t}}_{-} I \tag{6}$$

The positive force as characterized by the plus sign in Equation (6) says that, given the sales, the higher the lottery tax rate, the higher the tax revenue. On the other hand, we *expect* that a negative sign for the relation between lottery participation α and lottery tax rate τ, i.e., $\frac{\partial \alpha}{\partial t} < 0$.

[7] Finally, the third and the fourth column is the regression result of the following linear regression model. $S_{t,rollover} = \alpha_0 + \alpha_1 J_{t-1} + \epsilon_t$. "$J_{t-1}$" is the jackpot size rolled in from the $t-1$th issue. This regression is only applied to the sales in the rollover samples, $S_{t,rollover}$. Sales in the regular draw are not taken into account since the jackpot size must starts from 0 for all the regular draws. While the regression coefficient is

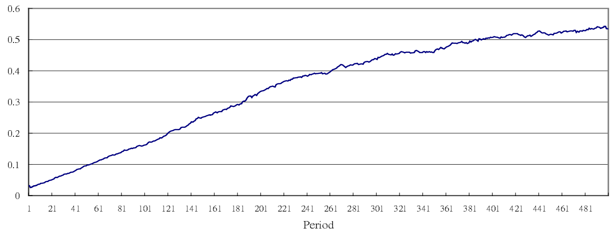

Figure 3: The Measure of the Belief of Fair-Game

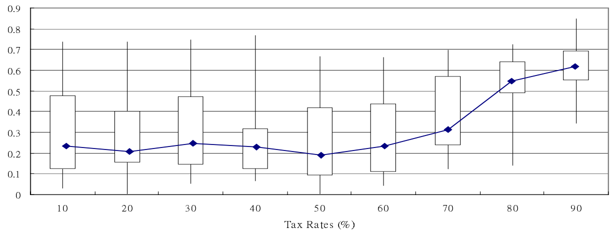

Figure 4: Value of the Regret Coefficient $\bar{\theta}$

5.3 Conscious Selection

Chen and Chie (2003) develop a metric to measure the degree of conscious selection, and by this metric one can have an idea of how the agent is far or close to a *fair-game believer*.

Figure 3 is the time-series plot of the d metric, which is derived by taking the average over all simulation samples (25 runs over different τs). As expected, d starts from a very low number due to random initialization, but then there is a tendency to a fair-game belief evolved since. Nevertheless, d does not converge enough to 1. Instead, it seems to settle around the level of 0.5 more so, which is approximately equivalent to a z of 14. Therefore, a degree of conscious-selection behavior is *weakly* observed. This result is almost the same as the one observed in Chen and Chie (2003).

5.4 Aversion to Regret

In this social learning framework, not only do agents learn from others, their preference may also be interdependent. By watching the evolution of the aversion to regret (characterized by θ), we can actually see how this interdependent preference may emerge or disappear. We examine the values of θ of all the 5,000 agents in the last period (period 500), and take an average from this sample. Call the average $\bar{\theta}$. Figure 4 is the box-whisker plot of $\bar{\theta}$ over the 25 runs. The line inside the box shows the median of the 25 runs. If we just focus on the median, we see that a relatively lower tendency to regret formed in these markets as opposed to that in [1] , while they share a similar pattern regarding to the impact of τ onto the level of θ.

6 Concluding Remarks

Agent-based modeling provides us a complete and systematic treatment of human behavior in complex adaptive systems. This paper provides an illustration of this idea. The various interesting behavior emerges from the bottom up allows us to see how each piece of this development actually

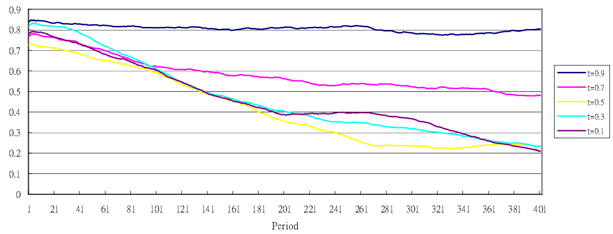

Figure 5: Rollover Frequency in different Tax Rates

support each other, and should better study them together in a coherent body, rather than treats them independently or exogenously. Thus the social euphoria extensively observed in this lottery market co-evolves with an increasingly active lottery participation , which in turn co-evolves with a declining rollover frequencies and rollover sizes (Figures 5) as well as a less interdependent preference (Figure 5).

Hence, agent-based modeling provides us also a tool to simulate evolution and learning, which is a typical model for bounded rationality. In addition to the co-evolutionary phenomena summarized above, evidence of learning is prevalent in this model, which includes belief updating and the associated decision on lottery participation , the emergence of the fair-game belief (Figure 3), and the less interdependent preference (Figure 5). All these figures shows strong tendency. The above-mentioned macro-dynamics is connected to these micro-behavior.

Having brought learning to our discussion, we are well alert of the lesson that *agent engineering matters*, or that learning or adaptive behavior can crucially change the final results. With this in mind, this paper consider a design which is different from [1]. This new design modifies the behavioral foundation of agents. Originally, it is based on fuzzy decision rules, and now is based on the expected-utility maximization framework. *Does this change matter?* The answer is largely *no*. Most results we have from [1] remain robust to this change. The phenomenon of the Laffer curve, the absence of the halo effect, and the emergence of the fair-game belief remains unchanged, at least, qualitatively. Even though the results remain unchanged, our understanding and interpretation of the same results may change because of different add-on behavioral mechanism. For example, the emergence of *social euphoria*, which is a main driving force to piece together most interesting simulated results, is simply not available in [1], when a model of subject belief is simply not there.

Bibliography

[1] Chen, S. -H. and B. -T. Chie (2003), "Agent-Based Modeling of Lottery Markets," *Proceedings of 7th Information Sciences (JCIS 2003)*, Cary, North Carolina, U.S.A., September 26-30, pp.1227-1230.

[2] Creight-Tyte, S. and L. Farrell (1998), "The Economics of the National Lottery," *Working Paper*, Department of Economics, University of Durham.

[3] Walker, I. and J. Young (2001), "An Economist's Guide to Lottery Design," *The Economic Journal*, 111, 700-722.

positive and is significant, its explanation power in terms of the coefficient of determination (R^2) is rather low as compared to the real data.