Lottery Sales and Income Distribution

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

Huei-Feng Fan
AI-ECON Research Center
Department of Economics
National Chengchi University
Taipei, Taiwan 11623
fan@aiecon.org

Abstract

In this paper, simulations based on the agent-based computational lottery market is applied to examine the relation between income distribution and lottery sales. Our initial results do not evidence that income distribution will impact on the lottery sales, and hence do not lend support to the relevance of income distribution to the lottery design. Furthermore, we also find that, regardless of the version of GA we used, lottery sales remain unchanged, i.e., the lottery sales performance does not depend on the assumed learning behavior, being its global (social) learning or local (community) learning.
Keywords: Halo Effect, Genetic Algorithms, Agent-Based Lottery Markets

Motivation

Lottery, as a way of fund-raising for public interests, has been implemented in more than 80 countries now.[1] Due to their possible differences in economic and social attributes, it is generally expected that, to achieve a certain goal, such as revenue maximization, different lottery designs are required. Nonetheless, at this moment, knowledge which can help us make such a differentiation is very limited. Before one can actually come up with such different designs, it is important to single out those contributing factors of the optimal lottery design. The purpose of this paper is then to examine whether the lottery design is sensitive to *income distribution*. We believe that this issue is relevant because the income distribution of countries who now implement lottery game are not homogeneous. Even for the same country, the income distribution would not remain unchanged at different phases of economic development,[2] which prompts us to question whether lottery designs should adapt to the change in income distribution.

Income Distribution Profile: The Case of Taiwan

In this paper, we simulate an agent-based lottery markets by assuming different income distribution profiles. These profiles are established by using the empirical income distribution in Taiwan from the year 1964 to 2003. The raw data are available from the Directorate General of Budget, Accounting and Statistics (DGBAS), Taiwan. For the consideration of smoothness, instead of using the short-run (annual) income distribution, we take a medium-term (five-year) income distribution by a simple average. The entire 40-year long data is divided into eight sub-periods, each consisting of five years. A simple average of the respective five years is then taken for each sub-period. The results are presented in Table 1.

Table 1 shows the average family income per household by *five* equal divisions of households according to personal income. For comparison purpose, numbers shown in this table has been standardized by assuming a fixed total income of these five families, which is set to 1,000. With these adjustment, for example, in the period 1999-2003, the highest 20% of families has an average income of 403.01, whereas the lowest 20% of has only 68.05. The wealthiest class is therefore 5.92 times higher the poorest class. The ratio of the wealthiest family class to the poorest family class is one kind of measure of income distribution. The last column of Table 1 indicates this ratio. Immediately right to this column is the well-known *Gini coefficient*.[3]From both

[1]Based on the *2002 World Almanac of Lottery*, 80% of the sales are concentrated in the Europe and North America.

[2]For the case of Taiwan, see Table 1 below.

[3]The Gini coefficient is a measurement of inequality developed by the Italian statistician Corrado Gini and published in his 1912 paper "Variabilita e mutabilita."

columns, we can see that the income distribution in Taiwan has been gradually become more disparate since middle 1970s. It would not be surprising to see this trend continuing for the next few years. It is, therefore, legitimate to ask what the impact of this changing income distribution to the optimal lottery design, optimal in the sense to maximize taxation revenue. To address the question, we may first propose the follow initial hypothesis: *income distribution exerts no effect on lottery designs.* If the hypothesis can not be rejected, then income distribution should not be a primary concern for lottery designers.

Having said that, we shall test our proposed hypothesis with an agent-based simulation approach by using the agent-based lottery markets proposed by Chen and Chie (2003). We shall fix a design and test the lottery market performance with different income distribution profiles. We first use the *uniform* income distribution as a benchmark to be compared with other income distribution profiles chosen from Table 1. We then compare these different scenarios under the same design and apply statistical tests to examine the significance of the observed differences. Furthermore, we also examine whether there is a functional relation between income distribution and lottery market behavior. Particularly, we are interested in knowing when income distribution become more disparate as the trend suggests, would the lottery revenue will positively or negatively affected? From a policy viewpoint, maybe the last question is more relevant.[4] Hence, four profiles of income distribution are chosen from Table 1 in a descending order of the Gini coefficients, and they are 1999-2003 (0.2590), 1964-1968 (0.2556), 1994-1998 (0.2469), 1989-1993 (0.2399).[5]

Table 1: The Average Family Income per Household by Five Equal Divisions of HouseHolds: Adjusted by a Fixed Size of 1000

Year	First 20%	Second 20%	Third 20%	Fourth 20%	Highest 20%	The High-Low Ratio	Gini Coefficient
1964-1968	77.75	123.95	164.35	221.75	412.20	5.30	0.2556
1969-1973	85.20	132.60	170.75	224.95	386.50	4.53	0.2317
1974-1978	88.98	135.81	173.32	225.03	376.86	4.23	0.2217
1979-1983	87.12	137.55	175.66	227.33	372.34	4.27	0.2201
1984-1988	82.30	135.45	175.21	228.13	378.91	4.60	0.2286
1989-1993	74.81	132.66	175.64	231.83	385.07	5.15	0.2399
1994-1998	72.33	129.37	174.55	232.87	390.89	5.40	0.2469
1999-2003	68.05	124.96	171.82	232.15	403.01	5.92	0.2590

[4]After all, uniform income distribution is never seen in any country of the world.

[5]Inside the bracket is the Gini coefficient.

Genetic Algorithms and Agent-Based Lottery Markets

Motivated by Chen and Chie (2003) and Chen, Chie and Lee (2005), we consider two different versions of agent-based lottery markets. These two versions differ with the formation of *lottery participation behavior*. In the first one, called the *Version 1*, the decision on lottery participation is formulated as a standard fuzzy inference, or, more precisely, the *Sugeno style of fuzzy inference*. The *antecedent* is a linguistic description of the jackpot size and is fuzzy, whereas the *consequent* is a numerical description of the lottery investment and is crisp. This modeling explicitly relates the participation decision to the jackpot size. The preference (utility) and subjective belief of winning is not involved. The second version differs from the former one by taking into account those fundamentals of agents, and modeling lottery decision within an *expected-utility* framework. The lottery participation is considered as a result of utility calculation.[6] Because of this difference, we also call agents in the first version, the *heuristic agents*, while agents in the second version, the *neo-classical agents*.[7] The purpose of that research is then to examine whether such a shape difference in agent engineering may actually result in any interestingly observed differences.[8]

When dispersion of income distribution is introduced to the model, the standard way of running genetic algorithms may be no longer applicable, because the link between strategies and its consequences is broken when some people are just born rich and some are just born poor. So, unlike the standard model (Chen and Chie, 2003), it is questionable whether the social-learning GA should be run over the entire population. Instead, an idea similar to the *island GA* may be applied, which is basically to run the GA only to each sub-population defined by the respective income class.

Nonetheless, whether we should run GA based on the entire population or just the sub-population defined by some innate attributes is more than just a technical question. In reality, it is possible that bounded-rational agents cannot correctly attribute the consequences to the causes. Under this situation, whatever rich people do is

[6] Details can be found in Chen, Chie and Lee (2005). The idea is as follows. We first derive the utility of non-gambling agents, and take this as a *benchmark*. We then assume the agent will buy the lottery ticket till the point that their expected utility is the same as the benchmark. We assume so because if considering lottery purchasing lottery itself is a fun activity, then such equivalence should actually make lottery participants be better off.

[7]"Neo-classical" in the sense that whether one should participate into lottery depends on whether it is better-off than not participating. See footnote 6.

[8]Interesting enough, the answer is largely *none*.

Table 2: Experimental Design

Market Parameters	
Pick x from X (x/X)	5/16
Lottery Tax Rate (τ)	50%
$s_0, s_1, ..., s_5$	0%, 0%, 35%, 15%, 12%, 38%
Drawing Periods ($\bar{r}$)	3
Number of Agents (N)	5000
Income (I)	Benchmark, D, E, F, G, H (Table 1)
GA Parameters	
Number of Fuzzy States	4 (for Ver1)
Number of Bits	4,4
Range of $p_{i,0}$	[0, 0.003] (for Ver2)
Periods (Generations) (T)	500
Crossover Rate (P_c)	90%
Mutation Rate (P_m)	0.1%
Tournament Size (φ)	200
Generation Gap (η)	100

always right and should be imitated. So, the main idea of running GA over the entire population is actually to assume a *weaker* rationality, which is not typical on the usual interpretation of GA as a social learning mechanism, but is not entirely illogical from the empirical observations of human behavior. Maybe, from a social scientist's viewpoint, it is interesting to run both version of GA for making a comparison. Therefore, we also consider both the standard GA and the island GA.

Experimental Designs

In a "x/X" lottery game, both a gambler and the lottery agency shall pick x numbers out of a total of X numbers, and then different prizes are set for different number matched. Let y denote the number matched. Clearly, $y = 0, 1, ..., x$. Let S_y be the *prize pool* reserved for the winners who matched y numbers. A special term is given to the largest pool, S_x, the *Jackpot*.

Each prize pool, S_y, shall be shared by all number of players who match y numbers, say N_y. The prize pool is defined by the *lottery tax rate*, τ, which is the proportion of sales that is not returned as prizes. Thus, the overall prize pool is $(1 - \tau)S$, where S is sales revenue and $1 - \tau$ is also called the *pay-out rate*. The overall prize pool will then be distributed to each separate pool based on a distribution $(s_0, ..., s_x : \sum_0^x s_y = 1)$, i.e., $S_y = s_y(1 - \tau)S$.[9] 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 2.

The market size is characterized by a number of 5,000 agents, each receiving a fix income every period. For the benchmark, the income is identically set to 200 for each agent. For the income profile D, E, F, G and H, income received by agents are heterogeneous, as detailed in Table 1.

In sum, we have a total of 24 designs, i.e., a combination of five income profiles, two versions of the agent-based lottery market, and two versions of the GA. To facilitate the aftermath statistical analysis, we conduct 50 independent runs for each design.

Experimental Results

The main focus is to see whether income distribution has impacts on lottery sales. In this simulation environment, lottery sales should be regarded as a random variable and its realization is different in each run, meaningful conclusions can only be made *statistically*. The least thing which one can test is to see whether there is any statistical difference in *mean*. But, in addition to mean, it is also imperative to know whether there is any difference in *distribution*. The second test is very useful for the stability concern.

References

Chen, S. -H. and B. -T. Chie (2003), "Agent-Based Modeling of Lottery Markets," *Proceedings of Agent2003-Conference on Challenges in Social Simulation*, Argonne National Laboratory and the University of Chicago, 225-233.

Chen, S. -H., B.-T. Chie, and C.-W. Lee (2005), "Lottery Markets–Design, Micro-Structure, and Macro-Behavior: An Agent-Based Computational Approach," *Proceedings of 4th Computational Intelligence in Economics and Finance(CIEF 2005)*,July 21-26, 2005, Marriott City Center, Salt Lake City, Utah, USA.

Directorate-General of Budget, Accounting and Statistics (2003), *The Survey of Income and Expenses in Republic of China, Taiwan (in Chinese)*.

La Fleur, Teresa (2002), *La Fleur's World Lottery Almanac*.

[9] In the event of $N_y = 0$, S_y is added to the next draw. A particular interesting case is $N_x = 0$. A common feature of lotteries is that, if there are no winners in a given draw, the jackpot prize pool from that draw is added to the pool for the next draw, referred as to a *rollover*. Rollovers usually enhance the attractiveness of the next draw, called the the *rollover draw*.