Lottery Markets-Design, Micro-Structure, and Macro-Behavior: An ACE Approach *

Shu-Heng Chen
AI-ECON Research Center
Department of Economics
National Chengchi University
Taipei, Taiwan 11623
TEL: 886-2-29387308
FAX: 886-2-29390344
E-mail: chchen@nccu.edu.tw

Bin-Tzong Chie
AI-ECON Research Center
Department of Economics
National Chengchi University
Taipei, Taiwan 11623
TEL: 886-2-29387308
FAX: 886-2-29390344
E-mail: chie@aiecon.org

Abstract

An agent-based computational modeling of the lottery market is established in this paper. Gamblers are modeled as autonomous agents with fuzzy logic and genetic algorithms. Three empirical observations from lottery markets are taken into account in the model of agents, namely, the halo effect and lottomnia, conscious selection of betting numbers, and aversion to regrets. The effect of the lottery tax rate (takeout rate) on tax revenue is studied in this framework. Our results show the existence of the Laffer curve. The Laffer curve indicates an optimal lottery tax rate at 40%, which is surprisingly close to the empirical tax rate averaged over the 25 lottery markets, 42%. The sensitivity of this result to the emergence of the interdependent preferences is also examined.

Keywords: Agent-Based Computational Modeling, Genetic Algorithms, Lottery, Halo Effect, Sugeno Fuzzy Models, Laffer Curve

1 Introduction

Agent-based computational modeling has become a very promising new research tool for economics. One of its main advantages is its encapsulation of the idea of *autonomous agents*. Through modern techniques of *agent engineering*, the researcher is endowed with rich expressive power of the *life* of agents. This rich expressive power not only helps us bridge the gap between the artificial world and the real world, but enables us to evaluate the consequences of some external interventions when the route from cause to effect becomes so complicated that it is hard to follow every step of it. Over the past decade, fresh and interesting insights have been brought to economic analysis in some active application areas of agent-based computational modeling, such as the artificial financial market. As an extension of our earlier studies in artificial stock market (Chen and Yeh, 2001; Chen and Yeh, 2002), this paper addresses an agent-based model of lottery markets.

*Research Support from NSC grant NO. NSC93-2415-H-004-005 is greatefully acknowledged.

Like the artificial stock markets, the research paradigm based on the representative agent already existed in the study of the lottery markets before the launch of agent-based modeling, such as Morgan and Vasche (1979), Morgan and Vasche (1982), Mikesell (1994), Mason et al. (1997), McConkey et al. (1987), Walker (1998), and Purfield and Waldron (1999). These earlier studies treated the demand for lottery tickets as an individual rational choice problem, and used demographic and socioeconomic data to estimate lottery demand. Nevertheless, our departure from the conventional research device to the agent-based modeling is motivated by the following two empirical observations.

Table 1: The Lottery Tax Rates

Nation	Official Issuer	Tax Rate	Commission Rate	Net Tax Rate
Austria	Austrian Lotteries	54.6%	9.30%	45.3%
Belgium	Lotterie Natlonale	48.4%	6.60%	41.8%
Brazil	Federal	68.4%	8.20%	60.2%
Canada	Loto-Quebec	48.7%	6.80%	41.9%
Canada	Ontario Lottery Corp.	51.2%	7.40%	43.8%
France	La Francaise	42.3%	5.00%	37.3%
Germany	Westdeutsche	53.0%	8.30%	44.7%
Italy	Lottmatica S.P.A.	48.4%	10.00%	38.4%
Italy	Sisal Sport Italia	65.4%	7.90%	57.5%
Japan	Dai-Ichi Kangyo Bank	54.2%	7.40%	46.8%
Spain	ONCE	50.4%	16.50%	33.9%
Sweden	Svenska Spel	48.8%	9.60%	39.2%
Taiwan	Taipei Bank	40.0%	8.40%	31.6%
UK	U.K. National Lottery	53.4%	5.10%	48.3%
USA	Ohio State	40.3%	6.40%	33.9%
USA	Michigan State	45.4%	7.00%	38.4%
USA	Georgia State	45.9%	7.00%	38.9%
USA	Maryland State	46.1%	5.70%	40.4%
USA	Illinois State	45.9%	5.10%	40.8%
USA	Texas State	46.4%	5.20%	41.2%
USA	New Jersey State	47.2%	5.40%	41.8%
USA	California State	49.3%	6.70%	42.6%
USA	New York State	49.4%	6.00%	43.4%
USA	Florida State	50.0%	5.60%	44.4%
USA	Pennsylvania State	49.1%	4.70%	44.4%

Firstly, Table 1 surveys *the lottery tax rate* of 25 lottery markets in the world. We see a quite wide distribution of the tax rate (the takeout rate).[1] From the lowest 40% in Taiwan to the highest 68.4% in Brazil, the difference is almost 30% high. Even in the U.S., there is a 10% gap from the lowest to the highest. The difference, which is also reflected in Figure 1, brings us closer to *the design issue*. But, the tax rate is only one

[1]The tax rate here refers to the gross tax rate, including what is reserved for bookmakers' commission. It is called the takeout rate, to be distinguished from the net tax rate. In this paper, the two terms, tax rate and takeout rate, will be used interchangeably.

dimension of the complex lottery design. Starting from the numbers offered to be selected, the matching rules, to the money to be awarded to different prize (such as the jackpot), one can face a great number of combinations (designs). Nevertheless, in literature, we see that little effort has been made to evaluate the impact of different designs, such as their effects upon lottery revenue.[2]

Since the lottery revenue is a major source for charity funds or education funds, it is imperative to have a mean to explore an extensive class of "what-if" scenarios. In this paper, the agent-based modeling, as an effective tool to deal with "what-if" scenarios, is used to analyze the effect of the tax rate on tax revenue. More specifically, we are interesting in knowing whether there is a *Laffer-curve phenomenon* in the lottery market. Stated slightly differently, is tax revenue globally sensitive to the tax rate? If so, what is the optimal tax rate? If not, within what range, it is insensitive, and is it wide enough to justify the empirical range shown in Figure 1?

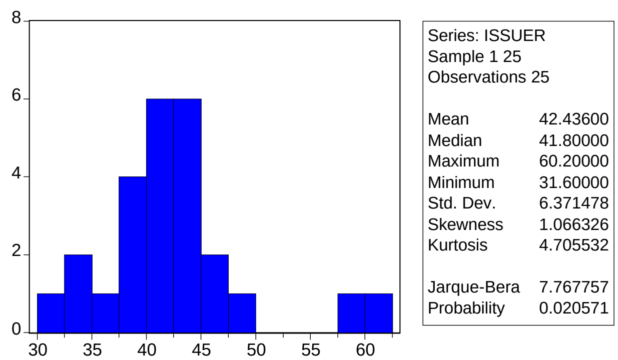

Figure 1: The Distribution of Tax Rates.

Date source: see Table 1.

The other empirical observation which motivates an agent-based model of lottery markets is the *psychology* of the lottery market. The ordinary gamblers seem to be not so much concerned with the probabilistic calculation of winning odds; instead,they rely on *heuristic strategies* for handling the available information. So, despite the fact that the series of the winning numbers is by all means generated by a random mechanism, they tend to believe that future predictions can be made on the basis of past history, and they tend to choose number in a non-random manner, called the *conscious selection*. There are even professional people who make a living by detecting "patterns". Griffiths and Wood (2002) provides a splendid review of various heuristics and biases involved in the psychology of lottery, such as the hindsight bias, representation bias (gambler's fallacy), the availability bias, etc. These heuristics and biases are, however, not easy to be captured by the standard rational analysis.[3] Nonetheless, in agent-based modeling, agents can be

[2]The only work known to us are Scoggins (1995), Hartley and Lanot (2000) and Paton, Siegel, and Vaughan (2002). The design of the United Kingdom National Lottery was not maximizing tax revenue was suggested by Hartley and Lanot (2000). Interesting enough, in October 2001, the U.K government implemented a dramatic shift in the taxation of gambling, and that results in a substantial decline in taxes levied on the U.K. bookmakers. An empirical study conducted by Paton, Siegel, and Vaughan (2002) indicated that the tax reduction caused a one-third reduction in duty receipts.

[3]A different but a related point is made by Farrel and Hartley (1998). They showed that repeated purchase of lottery tickets cannot be explained using expected utility functions. To regain the explanation power of the EU theory, it was assumed that agents are able to get some "*fun*" from gambling activity.

initialized with various cognitive considerations: the description and design of agents is basically open-ended.

In this paper, our agents shall be initialized with two heuristics and one psychological force. The first heuristic indicates agents' *portfolio strategies* (betting stake) based on their perception of the winning odds. The general observation that agents' betting momentum increases with the jackpot prize is what can come out of this heuristic. This heuristic, however, may have nothing to do with the sophisticated calculation of winning probability. In reality, the grand prize were generally well publicized, which creates an additional excitement referred as to *lottomania* (Beenstock, et al, 1999). Lottomania takes possession of the public, and attract their greater involvement.[4] The second heuristic reveals agents' *perception of the winning-numbers patterns*. No matter how fair or how random the winning lottery numbers were generated, gamblers tend to believe that some sequence of numbers are less likely than others. For example, a sequence of consecutive numbers, say, 1,2,...,6, are considered more improbable than other sequences. Finally, a psychological force added to agents' initialization is a *feeling of regret*, or known as *the aversion to regret* in the literature (Statman, 2002). Usually, when the mass media intensively report the winners with their gigantic prize, it may make those people who did not gamble feel regrettable: had they bet, the prize would have been theirs.[5] This psychological force referred as to the *regret effect* indicates that the interdependence of agents' utility function.

The two-above mentioned heuristics and psychology are first randomly generated to initialize agents' characteristics. These characteristics will evolve over time as agents are presumably utility-maximizers. As what has been popularized in the literature of agent-based economic models, the evolution will be driven by *genetic algorithms*.

The paper will be organized as follows. Section 2 will introduce an agent-based model of the lottery market. Section 3 will detail the use of the genetic algorithm. Section 4 outlines the experimental designs. The simulation results with discussion are given in Sections 5 and 6 respectively. Section 7 wraps up the paper with concluding remarks.

2 An Agent-Based Model of the Lottery Market

2.1 The Lottery Market and Its Design

Typically, an agent-based model comprises of two parts, namely, the *environment* and the *agent engineering*. The environment is characterized by a set of rules of the game, governing how agents are connected to the system and to other agents in the system. Here, it concerns a set of rules of a lottery game or a design of a lottery game.[6] Generally speaking, a lottery game can be parameterized by two parameters (x, X). 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*.

[4] In a way, this observation can be related to the *availability bias*, as initially proposed by Kahneman and Tversky (1973).

[5] Similarly, if the winner is absent, those who did not gamble may now have a degree of comfort as "I knew it".

[6] We would like to draw readers' attention to Walker and Young (2001) for an excellent introduction to the design of a lottery game. They also stimulated the issue of an optimal design.

Each prize pool, S_y, shall be shared by all number of players who match y numbers, say N_y. 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*. 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., $S_y = s_y(1-\tau)S$. 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 control panel of our agent-based lottery software (Figure 2).

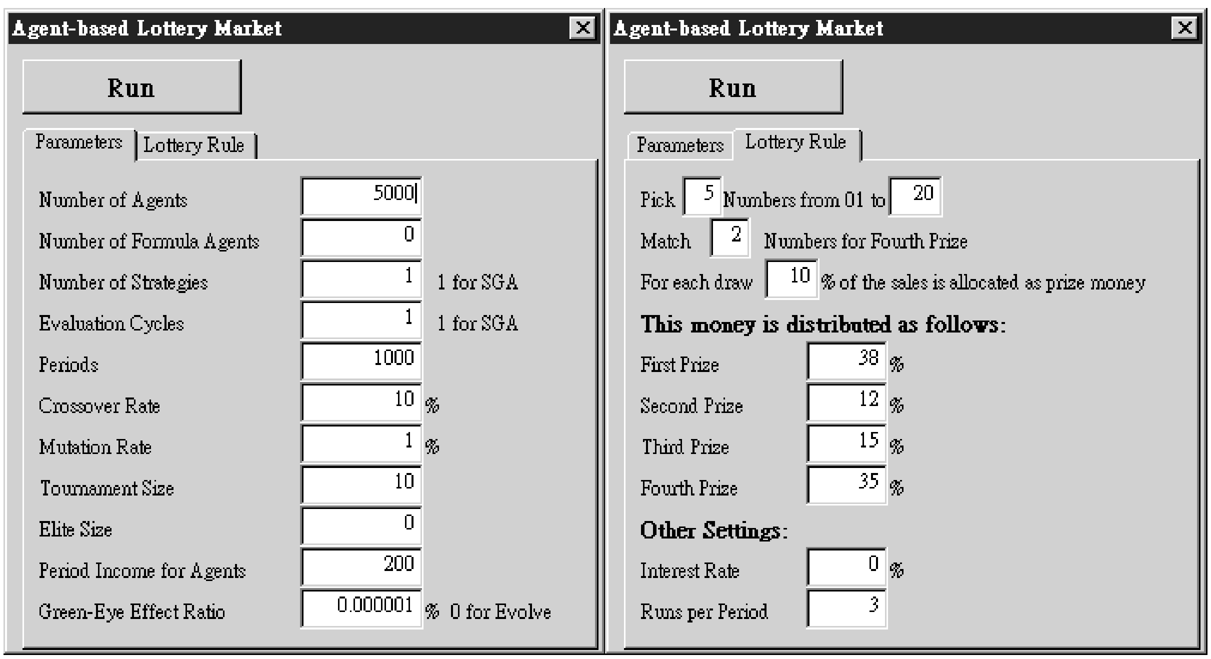

Figure 2: The Setting of Parameters and Environments

A purpose of agent-based simulation of the lottery market is to see how the changes in the design $\mathcal{L}$ can affect sales revenue, and more importantly, tax revenue. This brings us to another dimension of the lottery market, i.e., the *likely size of the market*. The likely size of the market can be determined by a series of economic and demographic factors. However, in this paper, we shall restrict our attention to only two factors, namely, *population size* and *income*. Both variables are treated as control variables in our agent-based lottery market. Let N denotes the number of agents in the market, and each of them are indexed by ι ($\iota = 1, 2, ..., N$). For simplicity, one can assume that their income y_ι is exogenously given and fixed. In the simplest case, y_ι is further assumed to be identical among all agents, $y_\iota = \bar{y}, \forall \iota$. $(N \times \bar{y})$ gives us only an upper limit of the market size. Lottery draws take place at regular intervals and, at each draw, agents decide how many tickets to purchase. Therefore, the actual market size is determined by *agents' participation*, which is the aggregation of the behavior of individual agents.

In literature, there are two approaches to analyze agents' participation in the lottery markets. The first approach is to use the empirical data to model the principal features of the observed aggregate behavior.[7] The second approach is to start from a rational model of *representative agents*, and then aggregate these representative agents.[8] The agent-based model is closer to the latter, while not using the devices of *rationality* and *homogeneity*. Agents are initially heterogeneous and boundedly rational, but they are autonomous and learning over time. Their details are left for the next subsection.

2.2 Agent Engineering

What do motivate agents to gamble, and how much to bet? We do not think that there is an unique answer or unique approach to this issue. Therefore, there are a number of possibilities in agent engineering. Nevertheless, a sound principle is to ground agent engineering with theoretical and empirical observations. Doing in this way, one can minimize the degree of arbitrariness. Our efforts to made in this agent-based model is to capture the following three "stylized facts" of the lottery market, namely,

- lottomania and the halo effect,
- conscious selection,
- aversion to regret.

2.2.1 Lottomania and the Halo Effect

That the lottery participation level is positively related to the size of the jackpot prize seems to be one of the most important empirical observations. The phenomenon that sales following a rollover are higher than sales prior to the rollover is known in the industry as the *halo effect* (Creigh-Tyte and Farrell, 1998; Walker and Yang, 2001). The halo effect is partially due to considerate media attention paid to rollovers, which in turn creates a bout of *lottomania*. Therefore, we can start building our agents from a *participation function* which relates the participation level to the jackpot size,

$$\mu = \rho(J), \tag{1}$$

where ρ is the participation function, μ is a measure of the participation level, and J is the size of the jackpot. The exact functional form of ρ depends on the framework within which the problem is formulated. In the standard rational analysis, μ is related to J via change in the *expected value*, or more generally, the *expected utility*, of the lottery ticket (Hartley and Lanot, 2000). However, here, we take a heuristic approach, and assume that gamblers base their decision on some heuristics rather than the possibly quite demanding work on expectations computation.[9]

[7]Papers belonging to this category are Farrell and Walker (1999), Farrell, Morgenroth, and Walker (1999).

[8]The number of papers in this category is much less. Hartley and Lanot (2000) is the only one known to us.

[9]Many details can complicate the computation of the expected value. First of all, the expected value depends on the expected number of the winning gamblers: the higher the expected number of winners, the lower the share of the jackpot for each winner. On the other hand, the expected number of winners depends positively on the participation level, by which the size of jackpot is also positively affected. This circular phenomenon applies to other non-fixed prize pools. Second, the expected value can differ among different agents, given their conscious-selection behavior.

By the heuristic approach, Equation (1) can be approximated by a few simple *if-then* rules. For example, "if the jackpot is unusually high, then I will spend 10% of my income to buy lottery tickets," or "if there is no rollover, I would spend only little." Notice that the antecedent or consequent of the rule contains the use of natural language which may not have concrete numerical meanings, such as the linguistic terms "high" and "only little" in the above example. While natural language has its ambiguities, people seem to be able to reason effectively with added vague and uncertain information and very often the decisions they make are the outcome of their approximate reasoning. Over the last four decades, we see the development of *fuzzy logic* as a formal approach to deal with these ambiguities. In this paper, we propose to represent the function ρ by a set of *fuzzy if-then* rules, which are manipulated by the standard mathematical operations of fuzzy sets as prescribed by fuzzy set theory.

We proceed as follows. First, let J_{t_r} be the jackpot prize updated at the rth day of the tth issue, where $r = 1, 2, ..., \bar{r}$. $\bar{r}$ denotes the gap between two draws. Suppose that the lottery draw takes place weekly, then $\bar{r} = 7$. Furthermore, let the set $\{J\}_{t_r}$ be the time series of the jackpot prize up to the time of t_r. Secondly, given the historical data, the *attractiveness* of the lottery game can be measured by how unusual of the J_{t_r} as compared to $\{J\}_{t_{(r-1)}}$, if $r > 1$, or $\{J\}_{(t-1)_{\bar{r}}}$, if $r = 1$. The agent will then act upon to the degree of attraction. For example, if the jackpot is "*huge*," the agent may react more energetically by betting greatly. Alternatively, if the jackpot is perceived as "*low*," they may be not interesting in spend a penny.

Technically, each agent gambles with his/her own *fuzzy rule-based system*, which comprises of a number of fuzzy if-then rules. Each fuzzy if-then rule of the system can be represented as follows,

$$If \quad J_{t_r} \quad is \quad A_i, \quad then \quad a_i. \tag{2}$$

A_i $(i = 1, ..., k)$ are *fuzzy sets* representing k different states of the jackpot prize. For example, consider the case $k = 4$, then A_1,...,A_4 can denote the following four linguistic descriptions of the size of the jackpot: "low", "medium', "high" and "huge." a_i is the level agent decide to participate given the current state is A_i. The participation level can be measured by the *proportion of income* which agents would spend to purchase lottery tickets. Call the vector $\vec{a}$ $(= (a_1, ..., a_k))$ the *participation vector*. Then different heuristics can be captured by different $\vec{a}$s. For example, $\vec{a} = (0.1\%, 1\%, 5\%, 10\%)$ characterizes the agent whose betting stakes is increasing in the size of the jackpot prize. On the other hand, $\vec{a} = (0.1\%, 1\%, 5\%, 1\%)$ indicates that the agent initially increases his betting stakes with the size of J, while when J is too large, he seems to lost the interest. The latter is particular suitable for the consideration that the expected number of winners increases with lottery sales, and hence higher sales imply *a small likely share* in the jackpot if the winning number is chosen.

Based on our description above, only the input set, A_i, of (2) is fuzzy, and the output set, a_i, is a crisp numerical value. This type of fuzzy rule is known as the Sugeno style of fuzzy rules, as distinguished from the Mamdani style of fuzzy rules, of which the input and output sets are both fuzzy.[10] Fuzzy sets are distinct from the classical sets (crisp sets) in the sense that the *membership* in the latter is *all or nothing*, whereas that in the former is a matter of *degree* (*more or less*). The degree is mathematically characterized

[10] It is, however, interesting to use Mamdani style of fuzzy rules by making the output set also fuzzy. So, for example, a heuristic can be "If J is *huge*, then play with *huge* stakes."

by a *membership function.* Formally, the fuzzy set A_i can be denoted as follows.

$$A_i = \begin{cases} \sum_{J \in N^+} \mu_A(J)/J, & if \;\; J \;\; is \;\; discrete, \\ \int_{J \in R^+} \mu_A(J)/J, & if \;\; J \;\; is \;\; continuous, \end{cases} \tag{3}$$

where $\mu_{A_i}(J)$ is the membership function, $\mu_{A_i} : N^+(R^+) \rightarrow [0,1]$.[11] The sign $\sum$ and $\int$ stand for the union of the membership grade. / stands for a marker and does not imply division.

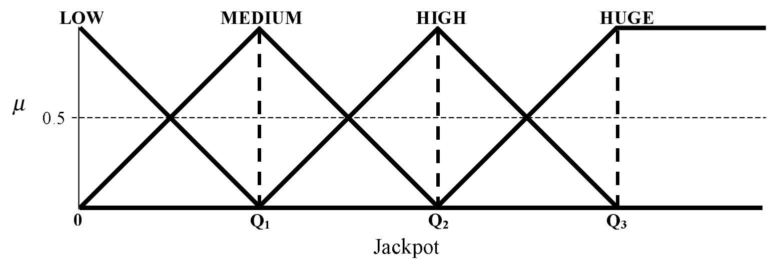

Figure 3: The Membership Function.

There are a wealth of membership functions. We, however, see little guidance as to the selection of them. Therefore, before more research been done on this area, we have to accept some degree of arbitrariness. For simplicity consideration, we choose the frequently-used *triangular-shaped fuzzy membership function.* Specifically, it is shown in Figure 3. In Figure 3, the domain of J is partitioned into four overlapping intervals by a sequence of base points Q_0, Q_1, Q_2, Q_3: $[Q_0, Q_1), (Q_0, Q_2), (Q_1, Q_3)$, and (Q_2, ∞). Denote them by $I_1, ..., I_4$ respectively. For each fuzzy set A_i, $\mu_{A_i}(J) > 0$ if $J \in I_i$; otherwise $\mu_{A_i}(J) = 0$. However, unlike the usual fuzzy membership functions, the base points upon which the membership functions are defined are *not fixed.* This is because all the linguistic terms have no absolute meaning. What perceived as *high* or *low* by agents will depend on what has happened before. It is the *frequency* which determines how we describe the event perceived. So, "*huge*" should refer to some events which happen more infrequently than what "*medium*" may refer to. This justifies the use of *sample statistics* as the base points, e.g., *quantiles.*

Since J will necessary start from zero, and will remain to be zero for all regular draws, only Q_0 can be fixed at 0. The rest of three base points can be taken as the first, second and third quantile of the sample $\{J\}_{t_r}/0$. Here, we only consider the quantiles after removing "0" from the time series sample $\{J\}_{t_r}$. What concerns us is as follows. Suppose the game is easy to win, then rollovers are infrequent and J_{t_1} has to start from zero for each draw. It may turns out that Q_1, Q_2,... are all zero, and the fuzzy membership functions shown in Figure 3 are, therefore, not appropriately defined.[12] The sample quantiles may converge if $\{J\}_{t_r}$ turns out to follow a stationary distribution; otherwise, they will change over time.[13]

[11] While in most lottery markets, J is a discrete integer, it is possible to have a design such that J can be a fraction or even real number. See Hartley and Lanot (2000).

[12] Take Taiwan national lottery as an example. The national lottery in Taiwan was introduced in Jan, 2002. Up to Feb. 25, 2003, it already had 127 draws. Out of these 127 draws, 111 were regular, and only 16 of them are rollovers. By these observations, Q_1, Q_2 and Q_3 are obviously all zero if 0 is left in the sample.

[13] If we further restrict the time-horizon of some agents' perception to be *finite*, then even though $\{J\}_{t_r}$

We now have assumed that all agents share the same membership function with the same way to determine Q(quantile)-statistics. As a corollary, their perception of all the linguistic states are identical. So, if the jackpot is perceived as "huge" for one agent in this draw, it is also perceived as "huge" for every other agents. Misunderstandings among agents on communicating these linguistic states are not allowed. Introducing heterogeneous membership functions can be computationally very demanding, and the necessity of doing so is not at all clear given the fact that human indeed are able to communicate very well with natural language.

The implementation of the fuzzy rules (2) proceeds as follows. For each period of time t_r, the agents observe the time series of the jackpot up to the beginning (the first second) of t_r, $\{J\}_{t_r}$. All Q statistics can be determined accordingly, and so is the membership function $\mu_A(J)/J$. Given J_{t_r}, i.e., the jackpot at the beginning of this period, the agent can then figure out the membership degree of each possible state (each fuzzy set), i.e., $\mu_{A_i}(J_{t_r})$ $(i = 1, ..., k)$. In the Sugeno fuzzy model, each corresponding rule is activated to a of degree $\mu_{A_i}(J_{t_r})$, and the output is a weighted average of all consequent actions a_i, weighted by the membership degree.

$$\alpha_{t_r} \equiv \alpha_{t_r}(J_{t_r}) = \sum_{i=1}^{k} \mu_{A_i}(J_{t_r}) a_i \tag{4}$$

Agents' involvement in lottery are defined by the fuzzy if-then rules (2) associated with the participation vectors $\vec{a}$. Adaptive behavior can be characterized by changes made in $\vec{a}$. In Section 3, we shall show how $\vec{a}$ can be encoded as a bit string and evolve via genetic algorithms.

2.2.2 Conscious Selection

The second important empirical observations of the lottery market is a general ignorance of the way probability operates. While all methods of selecting lottery numbers presumably have an equal probability, some of the general public do not seem to believe that the probability of the some numbers, say 1,2,3,4,5 and 6, being picked are as equally likely as any other sequence of six numbers. The phenomenon known as the *conscious selection* refers to non-random selections of the combinations of the numbers. What is even more interesting is that there is even a market for "experts" (like "technical analysts" in the stock market), who give advice to gamblers (investors) on which numbers to choose.

To take conscious selection into account, let $\vec{b}$ 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. Therefore, $\vec{b}$ shows a list of numbers which may be consciously selected by the agent. If $\vec{b}$ 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 $\vec{b}$ 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 $\vec{b}$ 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.

follows a stationary distribution, the respective finite-sample quantiles may still change over time. For these agents, what means to them by "high", for example, changes over time even though the distribution of J is stationary.

Agent's betting heuristic, h, is the fully characterized by the vector

$$h = (\vec{a}, \vec{b}).$$

To make it apparent that h are different over time (evolving) and are different over space (heterogeneity), we shall denote the heuristic used by the agent ι at time t by $h_{\iota,t}$. Section 3 shall detail the implementation of the evolution of $h_{\iota,t}$ via genetic algorithms.

2.2.3 Aversion to Regret

The last feature of our model of agents is the *utility function.* For simplicity reasons, most ACE models assume an exogenously-given utility function which is homogeneous among agents. However, we have slightly departed from this tradition mainly motivated by the following empirical observation, called the *aversion to regret.* Regret is the pain we feel when we find, too late, that a different choice would have led to a better outcome. In the case of the lottery market, regret simply means the utility of not to gamble depends on whether there are winners. If no body wins, that would make those who do not gamble feel no regret; however, if someone win, they may feel regret because it could be his had he given it a try. Lottery promoters capitalize on the aversion to regret when they encourage lottery buyers to keep on buying. If regret does play an important role, then agent's utility function is no longer independent.[14]

For simplicity, let us assume that agent ι has a simple one-period linear utility function of consumption:

$$u(c) = c, \tag{5}$$

with the budget constraint:

$$c \leq e - \alpha(\vec{a})e + \pi, \tag{6}$$

where e is his initial income, α is the proportion of his income spent on lottery, and π is the lottery prize. Since for those agents whose α is zero, their utility depends on whether there is a jackpot winners. The utility function (5) has to be modified as follows.

$$u(c) = \begin{cases} (1-\theta)c, & if \quad \alpha = 0 \quad and \quad N_x > 0, \\ c, & otherwise. \end{cases} \tag{7}$$

The θ in the utility function (7) measures how regretful the non-gambler would be if the jackpot is drawn.[15] 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 ($N_x = 0$). As a result, the utility function (7) can be extended as follows.

$$u(c) = \begin{cases} (1-\theta)c, & if \quad \alpha = 0 \quad and \quad N_x > 0, \\ (1+\theta)c, & if \quad \alpha = 0 \quad and \quad N_x = 0, \\ c, & otherwise. \end{cases} \tag{8}$$

[14] In spirit, this consideration is line with the regret theory proposed by Bell (1982) and Loomes and Sugden (1982). Regret theory offers explanations for numerous evident violations of the expected utility theory axioms. Regret theory says that the agent after making their decisions under uncertainty may have regret if their decisions turn out to be wrong even if they appeared correct with information available *ex ante*. This very intuitive assumption implies that agent's utility function among other things should depend on the realization of not chosen, and in this sense irrelevant, alternatives.

[15] Certainly, regret may work on the reverse direction as well. Nevertheless, since generally mass media will only give a large converge on the jackpot winners, and are not interested at all in anything happening to those non-gamblers, that asymmetric coverage makes the regret in the reverse direction rather negligible and hence it is assumed away in this paper.

Obviously, the larger the θ, the less independent of agent ι's utility. While we can treat θ as an exogenous variable, from a psychology viewpoint, it would be interesting to see how θ is determined endogenously. In this way, θ is treated as a personal *trait*, which indicates how agents experience things and his feeling about it. As a subject of development psychology, that trait can evolve over time. For example, the agent may learn to be more independent, and become not so much care about what others have; or, on the contrary, he may become less independent and would always like to compare what he has with others.[16] This evolutionary setting makes agents' preferences be also heterogeneous, adaptive and endogenously determined, a feature which is an unexploited potential of the ACE modeling.

To wrap it up, agents in our artificial lottery markets are fully characterized by the vector

$$(h_{\iota,t}, \theta_{\iota,t}) = (\vec{a}_{\iota,t}, \vec{b}_{\iota,t}, \theta_{\iota,t}), \tag{9}$$

where $\theta_{\iota,t}$ is the preference parameter of agent ι at time period t. The vector $(h_{\iota,t}, \theta_{\iota,t})$ will be encoded as a bit string, and then genetic algorithms is applied to evolve a population of $(h_{\iota,t}, \theta_{\iota,t})$, which is to be detailed in the next section.

3 Genetic Algorithms

3.1 Representation

Genetic algorithms (hereafter, GA) are motivated by natural genetics, and are developed by explicitly mimicking the process of natural genetics. A "*chromosome*" is the basic unit of GA. Like natural genetics, chromosomes in GA are strings of "*genes*", which define the characteristics of an individual. The way to represent a chromosome is called the *coding scheme*, and there are several coding schemes in GA. The most standard one is to use discrete values, such as binary, integer, or any other system with a discrete set of values. Among all the discrete-value coding system, binary coding is the most popular one. Binary coding represent each chromosome with a string binary digits, 0 or 1, also called the *bit string*. Chromosomes only gives the *genotype* of an individual, but the corresponding *phenotype* can be derived from a specified *decoding scheme*, also known as the *growth function*. In our model, binary coding is applied to the vector $(\vec{a}_{\iota,t}, \vec{b}_{\iota,t}, \theta_{\iota,t})$, which fully characterizes an individual ι at time t. However, since each component of the vector is associated with different function. The coding and decoding scheme would be different.

First, let us start with $\vec{a}$, the participation vector is a k-dimensional vector, $(a_1, ..., a_k)$, where a_i $(1 \leq i \leq k)$ lies between 0 and 1. Each a_i is first coded by a binary string with length l_a. The decoding is performed in the following way:

$$a = \frac{\sum_{i=1}^{l_a} c_i 2^{i-1}}{2^{l_a} - 1}, \tag{10}$$

where c_i is the cth bit counted from the right. So, totally, $\vec{a}$ is coded by a $k \cdot l_a$ bits, exemplified as follows.

$$\overbrace{(\underbrace{01..0}_{l_a \ bits} \mid \underbrace{10...1}_{l_a \ bits} \mid \underbrace{11...1}_{l_a \ bits})}^{k \ \ sub-strings}$$

[16]The snob effect or the bandwagon effect well taught in economics are just other examples demonstrating the interdependence of agents' preferences.

Figure 4 illustrates *a fuzzy inference system* (2) with $k = 4$ and the corresponding binary string of $\vec{a}$ (with $l_a = 4$), decoded as $\vec{a} = (0.2, 0.6, 0.8, 1.0)$. The input J is perceived by the agent, and the membership degree of each fuzzy set is calculated as follows: $[\mu_{A_1}(J), ..., \mu_{A_4}(J)] = [0, 0, 0.75, 0.25]$. So, by Equation (4), the agent will invest $\alpha = \sum_{i=1}^{4} \mu_{A_i}(J)a_i = 0.95$ of his income to purchase the lottery tickets.

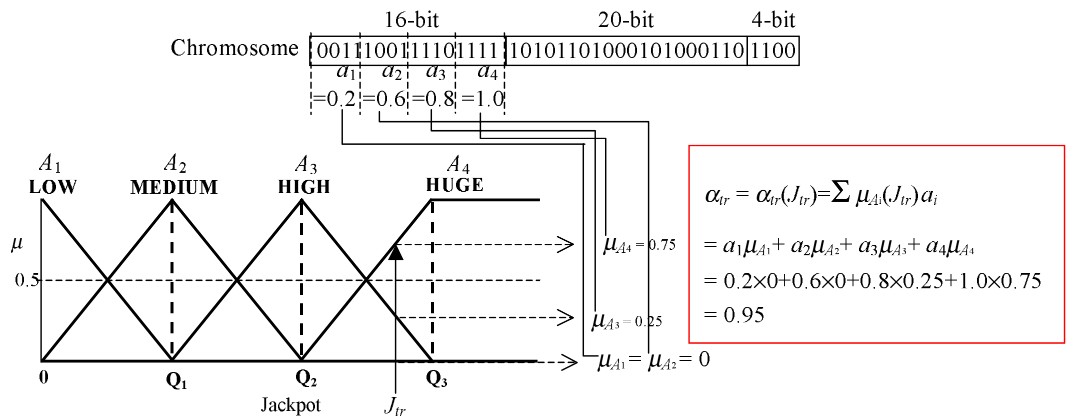

Figure 4: Betting Heuristics Based on the Sugeno Fuzzy Inference System

Second, is is straightforward to code the $\vec{b}$, the number-picking vector. As what we mentioned in Section 2.2.2, it is simply a X-bit string. An example of the case $X = 20$ is shown in Figure 5.

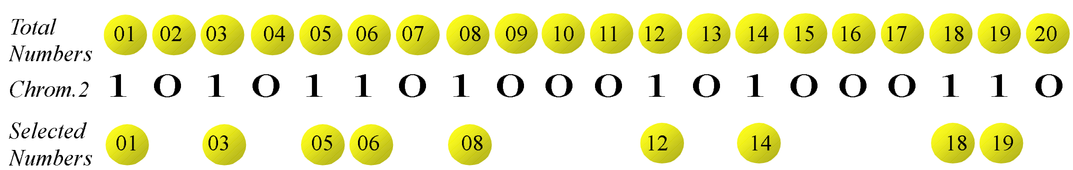

Figure 5: An Example of Agents' Picking Numbers

In Figure (5), the chromosome shows that the nine numbers 1, 3, 5, 6, 8, 12, 14, 18, 19, are consciously selected by agents. Consider a case of $x = 5$, i.e., each ticket can have only 5 numbers on it, then the five numbers shall be randomly selected from the total of C_5^9 (= 126) combinations.[17] If the numbers consciously selected by the agent is less than five, then the rest will be randomly pricked from those unassigned numbers.

Finally, the regret parameter θ, which also lies between zero and one, can be encoded in a similar fashion as Equation (10) by a l_θ-string bits. Therefore, the full characterization is encoded by a string with a total of $k \cdot l_a + 20 + l_\theta$ bits.

3.2 Evolutionary Cycle

Genetic algorithms starts with an initialization of a population of chromosomes (binary strings), called Generation 1 (GEN 1). The number of chromosomes or the *population*

[17] In the case of Taiwan lottery, the computer can automatically generate all 126 the (= C_5^9) tickets cover by the nine selected numbers. So, if the agent have enough budget (participation level), he will be able buy all of them. In practice, it is called the **combination strategies** (*bao-pai* in Chinese).

size, denoted by *Pop*, is fixed during the whole evolution. Then a fitness criterion (fitness function) is used to evaluate the performance of each chromosomes. Based on the performance evaluation, the next generation of chromosomes, shall be genetically produced by the incumbent. The genetic production starts from a selection of a *mating pool*. Usually the selection is biased toward the well-performed chromosomes to facilitate the implementation of *the survival of the fittest principle*.

There are two major selection schemes in GA, namely, *the roulette-wheel selection*, and the *tournament selection*. While these two selection schemes are well studied in the GA literature, which one is more suitable for agent-based economic modeling is still an open issue. This is because some advantages or disadvantages of these two schemes known to GA theorists may not be so much relevant for social sciences oriented studies.[18] Chen (1997) argued that, for social scientists, it is the *network* behind the social dynamics the primary criterion of the selection scheme. Roughly speaking, the roulette-wheel selection scheme implicitly assumes the existence of a well-connected global network, whereas the tournament selection only requires the function of local networks. Without further evidences on which network assumption is appropriate, it would be beneficial to try both selection schemes for the robustness concern. However, to have a better focus, only *tournament selection* will be tried in this paper. Nevertheless, according to the progress we make, the other selection scheme will be included at a later stage.

By tournament selection, each individual in the mating pool is determined as follows. We first randomly select φ random chromosomes without replacement, and then take the best *two* of it. The parameter φ is known as the *tournament size*, and it is also the mating-pool size. The two best out of the φ is called *parents* in GA. Two genetic alterations are operated on them to produce two offspring. The first alteration is *crossover*.

Crossover mainly exchanges genetic information of the two parents, and it is mainly through this process that the *exchange of information* among agents is done. What was exchanged (acquired and replaced) shows us, in a concrete way, *what agents had learned*. It is, therefore, crucial to represent the *target* which agents may cognitively recognize to learn. In the bit-string representation, there are two common ways to represent these targets: first, by *single bits* of the chromosome, and second, by *blocks* of bits. Corresponding to these two different representations, two crossover styles arise in the literature. *Uniform crossover* identifies the single bits as targets to learn, whereas *point crossover* emphasizes the role of blocks. While the latter has been widely used in economics, discussions of its decency over the former is not established yet. In this paper, we shall first follow the convention and use point crossover. However, the alternative will be kept in mind for the future robustness check.

Point crossover cuts each parent chromosomes into κ pieces. The cut-off points are randomly determined, but are kept in the same position between the parents. Then some of these pieces are exchanged to each other. It is called *one-point crossover* when κ is set to 1. Example of one-point crossover is shown in Figure 6.

As what has been shown in Figure 6, a cut-off point, which is same for both parents, is randomly selected. Then two offspring are formed after exchanging the parts being cut to each other. Since each chromosome represents altogether three different aspects of agents' behavior, the crossover operator is made in a *pair-by-pair* manner, i.e., to restrict the exchange only to the paired characteristic, called *paired crossover*. As what shown

[18]For example, from an optimization viewpoint, the roulette-wheel selection has two well-known disadvantages. First, it is *a danger of premature convergence* because outstanding chromosomes may take over the entire population very quickly. Second, the *low selection pressure* when fitness values are near each other.

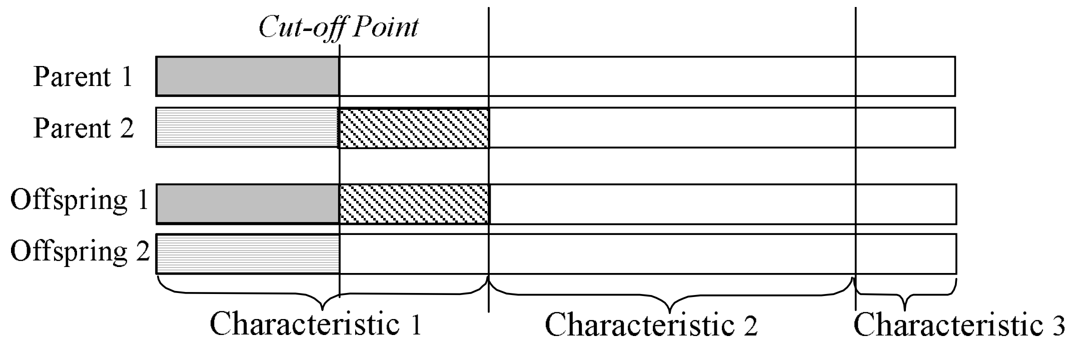

Figure 6: The Examples of One-Point Paired Crossover

in Figure 6, if the cut-off point is targeted at the first characteristic (betting strategy), then only the first characteristic will be cut and exchanged. The rest two characteristics remain unchanged.

Crossover can generate new genetic materials (new behavior), which may bring some desirable as well as undesirable features. Therefore, to be able manage or avoid the undesirable feature, a parameter called *crossover rate* (P_c) is introduced to GA. The crossover rate is the probability of turning on the crossover operator. If the crossover operator is turned on, the two parents will have two offspring as described above; otherwise, if it is turned off, then the result will be the two parents themselves. In a social learning framework, the crossover rate characterizes how quickly or easily agents will "learn" from others.[19]

The second genetic alteration is *mutation*. After crossover, each bit of the resultant chromosome has a chance of being flipped from "0" to "1", or "1" to "0". Ideally, the crossover operator enables agents to *discover* new types of characteristics by recombining the existing genetic materials, whereas the mutation operator make it possible for agents to discover something non-existing yet. The probability of being flipping is called the *mutation rate* (P_m), and it functions in a similar vein to the crossover rate. Nonetheless, the mutation operator can be more devastating; so the mutation rate is generally set very low.

The offspring after the mutation process will then to *replace* the old generation. There are two replacement strategies used in GA, namely, *generational replacement* and *steady-state replacement*. The former will complete replace the old generation, while the the latter will only replace the worst η per cent of the old generation. Clearly, the former is a special case of the latter when $\eta = 100$. Replacement is a consequence of *survival pressure*, which forces relatively ill-performed agents to change. By the parameter η, the agents belonging to the top $1 - \eta$ per cent would remain, and the agents belonging to the bottom η per cent would be replaced by offspring.[20] One may expect that the higher the survival pressure, the higher the η. While our software can allow users to choose the

[19] "Learn" is a positive word, but there is no guarantee that crossover will bring to a better result for agents. So, an alternative term to use is *"be influenced" from others*. To avoid the disturbing effect from the alteration of genetic information, Jasmania Arifovic invented the *election operator*, and applied to her series of studies. Some limitations of this operator have been discussed in Reiechmann (2002).

[20] The steady-state replacement is somewhat rigid, because it assumes that agents have a clear-cut decision on whether they should adjust their strategies, *yes* or *no*. Chen and Yeh (2001) introduce simulated annealing to make this originally crisp decision be a probabilistic one.

η, to focus better, this paper will only consider the case $\eta = 100$, which is equivalent to the generation replacement. So, even the best of the old generation (the elite) will not be kept.

We wrap up the section with a pseudo program shown in Figure 7. We begin with generation 1 ($Gen = 1$), initialize the first generation of population (POP_1), an then evaluate it. After that the evolution cycle starts by iteratively following the step of selection and genetic alterations until Pop offspring are generated. These offspring (POP_2) will then replace the old generation (POP_1), and we move to Generation 2, followed by evaluation, selection, genetic alterations, replacement, and moving to Generation 3. This cycle is going on and on until the termination criterion, which is the number of generations to evolve, is met.

The control parameter discussed in this section can all be input directly from the control panel of the software (see the left panel of Figure 2).

```
begin
    Gen := 1;
    Pop := Population-Size;
    initialize(POP(Gen, Pop));
    evaluate(POP(Gen, Pop));
    while not terminate do
    begin
        for i := 1 to Pop step 2
            Parent1 := Tournament-Select-1st(POP(Gen, Pop));
            Parent2 := Tournament-Select-2nd(POP(Gen, Pop));
            OffspringPOP(Gen, i) := Crossover-Mutation-1st(Parent1, Parent2);
            OffspringPOP(Gen, i+1) := Crossover-Mutation-2nd(Parent1, Parent2);
        next i
        evaluate(OffspringPOP(Gen, Pop));
        POP(Gen+1) := OffspringPOP(Gen, Pop);
        Gen := Gen+1;
    end
end
```

Figure 7: The Pseudo Program

4 Experimental Designs

The agent-based lottery market as introduced in Sections 2 and 3 is summarized by two sets of parameters, the one associated with the *market*, and the one associated with the *agents*. Parameters associated with the market are encapsulated into the vector $\mathcal{M}$.

$$\mathcal{M} = (x, X, \tau, s_0, ..., s_x, \bar{r}, N, \bar{y}).$$

Except for the last two parameters, N and $\bar{y}$, which is beyond the control of the lottery administration, the rest parameters are just part of the design of the lottery game. As far as the optimal design is concerned, one may be interested in knowing whether there is a set of parameter values,

$$(x^*, X^*, \tau^*, s_0^*, ..., s_x^*, \bar{r}^*),$$

such that, given N and $\bar{y}$, an objective function can be optimized, e.g., the tex revenue can be maximized.[21] While this is the most general way one can pose the design issue, it does not provide us an interesting focus. Being an initial stage of this research line, this paper chooses τ as the first direction to move. As what we have briefly surveyed in Section 1, the lottery tax rate is indeed a point of attraction, not just from the academic perspective, but more also from the public-policy perspective.

This paper studies the possible relation between the lottery tax rate and the tax revenue by hypothesizing the existence of a Laffer curve, and hence an optimal interior τ. To do so, different values of τ ranging from 0 to 90% are attempted in this paper. The rest of the market parameters are treated as *constants* throughout the entire simulation, and they are listed in Table 2.

Table 2: Experimental Design

Market Parameters	
Pick x from X (x/X)	5/16
Lottery Tax Rate (τ)	0%, 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 (y)	200
GA Parameters	
Number of Fuzzy States (k)	4
Number of Bits (l_a, l_o)	4,4
Periods (Generations) (T)	500
Crossover Rate	90%
Mutation Rate	0.1%
Tournament Size (φ)	200
Generation Gap (η)	100

Since we are using the current version of the ACE model only as a *toolbox* for enhancing our understanding of the lottery behavior, these parameters are not chosen through *calibrating*. Nonetheless, the set of prize ratios, $s_0, ..., s_5$, are chosen to be consistent with the National Taiwan Lottery. Similarly, the drawing periods of each issue ($\bar{r}$) is also motivated by it, assuming that each period is equivalent to one day, and there are two issues per week. The most intriguing part, however is the setting of x/X. Obviously, the choice of x and X affects the likely number of winners. If $x = 6$ and $X = 49$ then the probability of a ticket being the winning combination is approximately 1 in 14 million, and if $x = 6$ and $X = 53$ then the chance of buying the winning combination is approximately 1 in 23 million.[22]

Based on this winning probability, the expected number of jackpot winners and their prize depends on the number of tickets sold, which in turn depends on the size of the market, i.e., $N \times y$. Given the size, a game design which is sensible for a country A is likely to be hard for country B, whose population size is relatively smaller, or people are

[21] In fact, one advantage of the ACE modeling is to make this complex design issue at least computationally solvable.

[22] These figures are directly borrowed from Walker and Young (2001).

less wealthy. Therefore, x/X cannot be set independently of N and y. In ACE model, we have a severe restriction to N. To run the simulation in a reasonable fast way, N can only be set as a number like from 5,000 to 10,000, which can hardly match the population size of a real country. This forces us to modify x/X in a way such that it can be comparable to a real market, say Taiwan. This makes us to consider a rather smaller X, which is only sixteen. A game of 5/16 is then matched to a market size $N \times y = 5000 \times 200$.[23]

The second set of parameters concerns the control parameters of the genetic algorithm.

$$\mathcal{A} = (k, l_a, l_\theta, T, \varphi, P_c, P_m, \varphi, \eta)$$

To have a focus and make our presentation easier, all these parameters are also fixed during the entire simulation, as given on Table 2. A remark is made to the parameter φ, i.e., the tournament size. This unusually large tournament size (φ=200) allows for a greater extent of interaction among gamblers, which is to approximate the intensive attention drawn to lottery results reported by mass media. We do admit the potential significance of some variations of $\mathcal{A}$. However, the sensitivity analysis will be important only after we are able to show something interesting with the current fixed design.[24]

Models built upon genetic algorithms are stochastic model in the sense that even the same fixed design may come up with different results. Therefore, to enhance the validity what we may conclude from the simulation, multiple runs of the same design is inevitable. Each set of parameters is run 25 times. Depending on the market participation level, the running time varied from 3 hours to 20 hours under Pentium 4 2.4G Hz.[25]

5 Experimental Results

5.1 The Take-out Rate and Tax Revenue

As declared on Table 2, each run lasted for 500 periods, i.e., 500 draws. Tax collected from each game is indexed by R_t (the tax revenue at the tth issue of the tickets). A time series $\{R_t\}_{t=1}^{500}$ is observed after each run. To make sense of the results, we further normalize the revenue series by dividing R_t by the total income $N \times \bar{y}$, and call this new series $\{r_t\}_{t=1}^{500}$ the *normalized tax revenue* series. Notice that normalized tax revenue can be interpreted as an *effective tax rate*. By convention we took away the first 100 periods of the data, and calculated the *mean* for the rest of sample, i.e., $\{r_t\}_{t=101}^{500}$. Denote it by $\bar{r}$. Since we have 25 runs for each single lottery tax rate, we, therefore, report the median of $\bar{r}$ over these 25 runs, and the results are shown in Figures 8.[26]

[23]What mainly concerns us is the expected number of jackpot winners. The expectation is based on the assumption that all income are spent on the lottery market. By that assumption, the expected jackpot winners are 229 in the artificial market, whereas in the real (Taiwan) market, it would be 381. The difference seems to be acceptable considering that in the real market there are many other tools for play gambling, while in our artificial market there is only one way to do so. But, by no means, we are doing any calibration in a serious way. Instead, we just want to make sure that the game running in our artificial market is not incredible easy or hard from a real-market perspective.

[24]In the future, we do plan to take this agent-based lottery market to some sensitivity issues pertaining to the choice of different selection schemes, market sizes, crossover style,...etc, including their economics significance and the effect upon the simulation results.

[25]Usually when the lottery tax rate is low, the associated participation level can be quite intensive, and hence computational load is very heavy.

[26]The reason that we report the median instead of the mean is because that the former is a robust statistics, whereas the latter is sensitive to outliers.

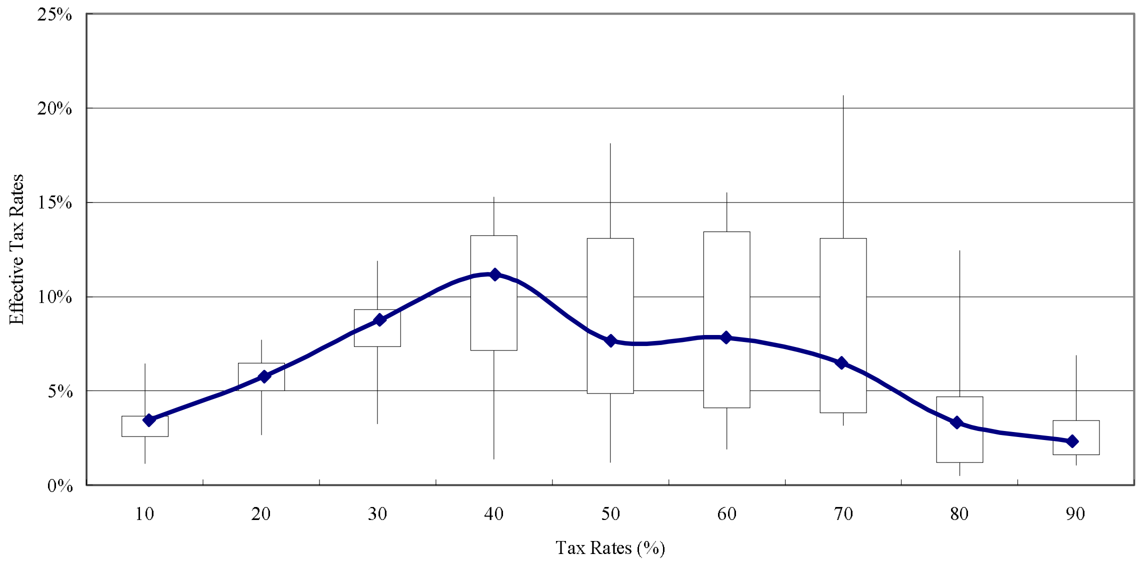

Figure 8: Tax Revenue Curve and the Associated Box-Whisker Plot

The figure shows that the (normalized) tax revenue first increases with the lottery tax rate τ, and then decreases with it. The highest tax revenue appears at $\tau = 40\%$ with a $\bar{r}$ of 10.5%. In addition to median, it is also interesting to notice the change in the uncertainty of tax revenue under different tax rates. This is reflected by the associated box-whisker plot also shown in Figure 8. The box in the middle of the plot covers 50% of the simulated tax revenue. The longer the box, the more uncertain the tax revenue. From Figure 8, the tax revenue is relatively low and stable when the tax rate comes to its two extremes ($\tau = 10\%, 90\%$). However, the box starts to inflate when when the tax rate is away from the two extremes, which signifies the growing uncertainty in tax revenue. The degree of uncertainty is further compounded by the enlarging whiskers, which extend the box to the frontier of the sample distribution.

Another way to describe what found in Figure 8 is that the *elasticity* of changes in tax revenues with respect to changes in the tax rate is *unstable*. Tax revenue can be *statistically insensitive* to a range of tax rate, say from $\tau = 0.4$ to $\tau = 0.7$. This makes the determination of the optimum tax rate τ^* less certain. Given this circumstance, a rigorous statistical approach must be taken to deal with this policy issue. However, in real world, it is very difficult to do so because the lottery administration cannot fine-tune the tax rate too frequently. As a result, it is difficult to get enough observations to support an estimation job. *This limitation shows the potential value of agent-based modeling in policy analysis.*

If one temporary accept the peak of the simulated Laffer curve as the optimum tax rate, i.e., $\tau^* = 40\%$, *is this number interesting*? In Figure 1, we have a survey over 25 lottery markets, the average lottery tax rate is 41.8 as the median, and 42.3 as the mean. They are *surprisingly* close to our $\tau^* = 40\%$. They are surprisingly close because, at this stage of research, we have not played around with our parameters to fit our results to the real data. The result is cooked up with only the three empirical properties pertaining to individual agents as discussed in Section 2.2. All aggregate results emerge from the interacting autonomous agents whose three imposed-characteristics evolved over time. These setups, while still simple, are hard enough to forecast what the aggregate results

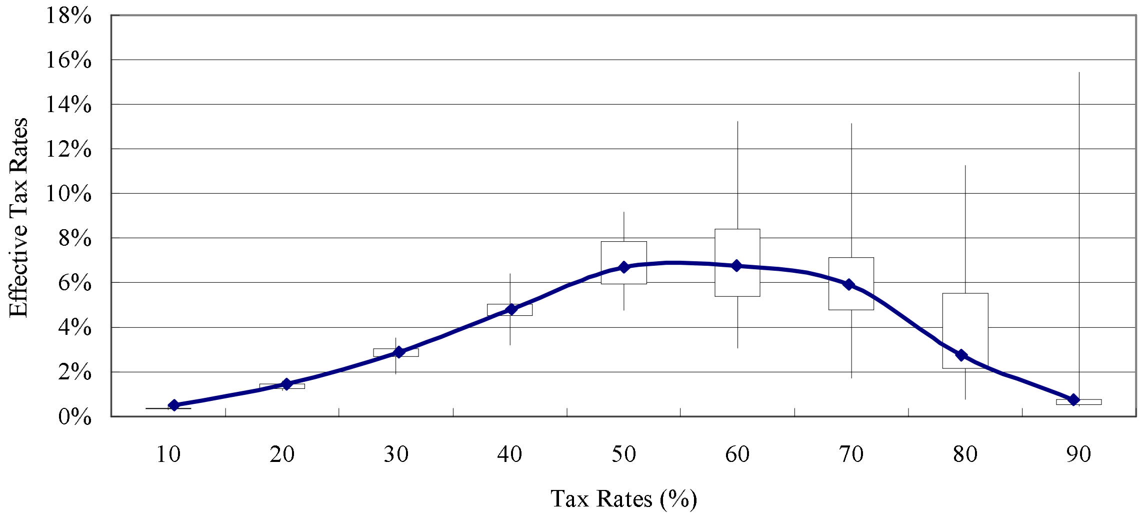

Figure 9: Tax Revenue Curve and the Associated Box-Whisker Plot under No Regret Effect

we may come up with, including a τ^* of 40%.

Certainly, that does not mean that the complex adaptive system like this one can have only one unique solution: $\tau^* = 0.4$. Is it possible that different settings of the parameter values can lead to different results? Or are we, somehow, just by luck simulate a system with a set of parameters whose optimal solution is consistent with the empirical observation? This is indeed the *robustness issue* that one have to face in ACE modeling. To be honest, there is no end to a full robustness check, as one can always go on and on. A compromise must be made somewhere, but that can made in an academically acceptable way. First, one have to make their software publicly available so that it can be always open to further validation, which we shall do so as what we did in the past.[27] Second, one then can restrict their examination to some non-trivial variations either in the part of agent engineering or the part of environment. At this point, this paper takes one interesting aspect to examine, i.e., *the regret effect. What does it happen if the regret effect is assumed away*?

Section 2.2.3 introduces the regret effect. It is introduced to the model via the parameter θ in Equation (8). θ is originally randomly generated between 0 and 0.93, and is heterogeneous among agents. It then evolves over time. Now, to assume away the regret effect, we simply fix θ to 0. We then re-run the experiments as designed on Table 2, and, as before, 25 runs for each τ. The results of the effective tax revenue and the associated box-whisker plot is drawn on Figures 9. Figure 9 shows an uni-modular Laffer curve which peaks at τ=0.6 with a $\bar{r}$ of 7%. As compared to Figure 8, the new Laffer curve levels down and the peak also moves to the right, which indicates that the regret factor may not only affect the optimal tax rate τ^*, but also the tax revenue.

That the absence of the regret factor can impact the optimal tax rate and tax revenue is an interesting observation. It has something to do with the endogenous evolution of the dependent utility function. We shall come back to this issue on Section 5.4.

[27] See the website: http://aiecon.org/software.htm.

5.2 Rollovers and Sales

In addition to the Laffer curve, a number of interesting observations of the lottery behavior can also be made from our agent-based simulations. We proposed three analyses in this paper. The ones to be discussed in the next two sections are behavior of *conscious selection of winning numbers* and *interdependent preference*, and in this section we address the empirical relation between *rollovers* and *sales*.

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. For example, based on the time series data of the UK lottery from November 19, 1994 to March 5, 2003, a total of 751 draws, the average sales is 56.0 million pounds over the rollover draws, whereas it is 41.4 millions over the regular draws. Nevertheless, exceptions exists. Among a total of 112 rollover draws of the UK lottery, it happened 25 times that sales actually *fell*.

To have a general picture of the empirical relation between rollovers and sales, Table 3 summarizes some basic statistics of three lottery markets: UK, Taiwan, South Africa, Ireland, Swiss, Japan, and Turkey. We first conduct a statistical test for the significance of the difference between the sales in the rollover draw and the sales in the regular draw. The t test statistics are shown in the second column. Below each test statistics are the corresponding p values.[28] Second, for those rollover draws, we further regress sales against the jackpot size as

$$S_{t,rollover} = \alpha_0 + \alpha_1 J_{t-1} + \epsilon_t. \tag{11}$$

"J_{t-1}" is the jackpot size rolled in from the $t-1$th issue. Regression (11) 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. The values of the coefficient α_1 and $\bar{R}^2$ are reported in Columns 3 and 4. Finally, as mentioned earlier, a surprise is to see that sales may fall in some rollover draws. To acknowledge the occurrence of this anomalous relation, the fifth column gives the percentage of the rollover draws whose sales actually declined rather than rose. We consider this statistics important because it hints that the underlying agents' behavior, which connects rollovers to sales, may be more complicated than one may hypothesize from a simple linear regression.

Table 3 shows quite consistent patterns for the seven lottery markets. First, the *halo effect* is evident in all markets. This is reflected by the significant positive t statistic (the second column), which means that sales in the rollover draw are significantly greater than those in the regular draw. Second, as we expect, the jackpot size significantly prompts sales. Its positive effect on sales is statistically significant in all markets. The only question is whether its explanatory power is good enough. In some markets, $\bar{R}^2$ is surprisingly high up to 90 per cents, whereas in the other two countries, it is only 20 per cents. However, what should not be hidden from this general expected results is the existence of anomalies. The anomalous relation between rollovers and sales are prevalent in the two of the three markets. In South Africa, sales fell in 28 per cents of the rollover draws, whereas in the U.K. it declined in 22 per cent of them. What may cause these anomalies is an issue which we would like to pursue in this line of study.[29]

Based on these references, it is interesting to see whether the similar patterns hold for our artificial lottery markets. Therefore, we do the same statistics over the simulated data. What we do here is to *pool* together all the simulated date under the same tax rate, and do the statistics by the tax rate. The result is shown on Table 4. Marked contrasts

[28] Here, a two-tail test is applied.

[29] Interesting enough, we have not seen any literature giving enough attention to this anomaly.

Table 3: Rollovers and Sales: Statistics from the Real Data

Nation/Period (periods)	t statistic (p-value)	α_1 (p-value)	$\bar{R}^2$	anomalies
United Kingdom (11/19/94-3/8/03)	4.2231 (0.0000)	1.9418 (0.0000)	0.2126	22.32%
Taiwan (1/22/02-5/23/03)	3.4578 (0.0013)	10.1766 (0.0000)	0.9133	0.00%
South Africa (3/11/00-6/4/03)	3.6959 (0.0001)	1.3307 (0.0000)	0.2052	27.78%
Ireland (1/2/02-6/28/03)	5.9171 (0.0000)	0.7451 (0.0000)	0.6371	1.04%
Swiss (1/1/03-6/28/03)	5.7976 (0.0006)	9.3881 (0.0015)	0.9227	0.00%
Japan (7/4/02-6/26/03)	5.9727 (0.0000)	1.1291 (0.0000)	0.9460	6.67%
Turkey (6/20/01-7/2/03)	0.5384 (0.3067)	12.6762 (0.0713)	0.6196	0.00%

between Tables 3 and 4 are observed. First, *the halo effect disappears.* More than that, all t statistics become now significantly *negative*. We now have the opposite of the halo effect, the *anti-halo effect*. Second, the effect of the jackpot size is by and large positive, which is consistent with what that observed in the real data. However, its explanatory power diminishes very fast with the increase in the takeout rate. Given the result above, it is not surprise to see that "anomalies" now become normal. For all takeout rates, sales declined in more than 50 per cents of the rollover draws.

The disappearance of the halo effect and the appearance of the anti-halo effect is certainly astonishing. This is even more so because our agent engineering is based upon the consideration of the halo effect (see Section 2.2.1). However, to compare what we have from the real data with what we have from the artificial date provides us a chance to reflect upon something which we may take it for granted. In particular, what is the essence of the phenomenon of the halo-effect? Why did the agent-based system built upon GA fail to deliver this feature? Also, given the halo effect, why are there so many exceptions (about 20% to 30% in real markets)? Why did the agent-based model is particular good at producing these "anomalies"? These are the questions to be addressed in Section 6.

5.3 Conscious Selection

Hard empirical statistics on conscious-selection behavior is not available yet in the real market, while the patterns of the lottery numbers have been analyzed by many "experts" who advise people how to pick the numbers. In our simulation, the numbers favored by each agent is observable. The vector $\vec{b}$, as detailed in Section 2.2.2, shows the numbers picked or excluded by the agents. This profile provide us the chance to observe the behavior of conscious selection. In particular, it enable us to address the question as to whether the agent essentially believe that winning numbers are randomly selected.

This can be done by asking each agents the following question. Does the agent believe that each number are equally likely (or unlikely) to be picked by the lottery adminis-

Table 4: Rollovers and Sales: Statistics from the Simulated Data

Tax Rates	t statistic (p-value)	α_1 (p-value)	$\bar{R}^2$	anomalies
0	-19.3379 (0.0000)	0.6014 (0.0000)	0.1352	49.14%
0.1	-23.0334 (0.0000)	0.6438 (0.0000)	0.2996	58.67%
0.2	-66.1523 (0.0000)	0.5583 (0.0000)	0.0645	63.63%
0.3	-99.0913 (0.0000)	0.1093 (0.0240)	0.0042	63.50%
0.4	-117.1700 (0.0000)	0.2144 (0.0000)	0.0148	62.09%
0.5	-100.7600 (0.0000)	0.1563 (0.0000)	0.0093	63.36%
0.6	-87.8737 (0.0000)	0.0322 (0.4165)	-0.0001	61.58%
0.7	-82.4286 (0.0000)	0.1121 (0.0462)	0.0010	60.52%
0.8	-49.3922 (0.0000)	0.0899 (0.0840)	0.0004	57.47%
0.9	-44.7909 (0.0000)	-0.2789 (0.0130)	0.0010	56.34%

tration? If the agent believes that winning numbers are randomly generated, then all combinations are available for him to select. Therefore, simply by counting how many combinations are excluded by the agent or how many combinations are effectively available for the agent, one can develop a metric to measure how far the agent is away from the belief of a fail game. Let d be the metric, and

$$d = \begin{cases} \binom{X-z}{x-z} / \binom{X}{x}, & if \quad z \leq x, \\ \binom{z}{x} / \binom{X}{x}, & if \quad z > x, \end{cases} \tag{12}$$

where z is the number of 1s appearing in $\vec{b}$.

When the agent believes that the game is fair and treats all the numbers equally, then $z= X$ (or 0), and the measure d achieves its maximum d_{max}:

$$d_{max} = \binom{X}{x} / \binom{X}{x} = 1. \tag{13}$$

On the other hand, if the agent has exactly x numbers in his mind, then the game for him is completely deterministic, and d gets to its minimum d_{min}:

$$d_{min} = 1/\binom{X}{x} \approx 0. \tag{14}$$

So, simply by watching how d is close to 1 or 0, one can have an idea of how the agent is far or close to a fair-game believer. A time series display of the metric shall shed light on how well the behavior of conscious selection is developed.

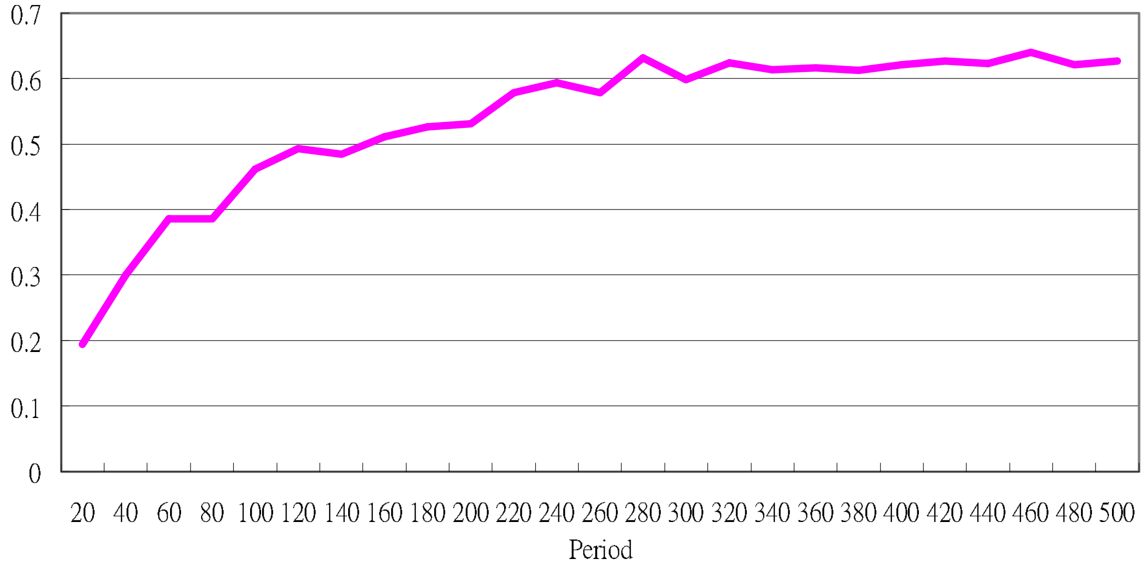

Figure 10: The Measure of the Belief of Fair-Game

Figure 10 displays the evolution of the metric d at a highly aggregation level. What shown in the x-axis is time. An observation is taken for every 20 periods. For each sampling period, we pool together the d of all 5,000 agents over 25 runs under all the tax rates. So, each $\bar{d}$ shown here is the average of $5,000 \times 25 \times 10$ individuals' d. The time series plot of $\bar{d}$ basically shows a monotone increasing behavior, which characterizes the gradual convergence to the belief of fair game. However, it does not converge enough to 1. Instead, it seems to settle around the level of 0.6, which is approximately equivalent to a z of 14. Therefore, a degree of conscious-selection behavior is weakly observed.

5.4 Aversion to Regret

As mentioned in Section 2.2.3, a positive θ in Equation (8) intensifies agents suffering' when they don't buy the lottery ticket, while later on someone takes away the jackpot prize. Nonetheless, that intensity is treated symmetrically so that extra comforts can be gained when there is no match. Given this potential, the agent can learn and develop an independent preference with $\theta = 0$, or to develop a less independent preference with a positive θ. It is interesting to see how a culture of θ is formed over the course of evolution and its implications.

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 11 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 tendency to regret is cultivated in the market. Given the dual role of θ, a more appropriate interpretation is that a culture which people are sensitive to what other people encounter is nursed in this lottery environment. In all cases, the median of $\bar{\theta}$ is high up to 0.8, and even to 0.9 with increase in the takeout rate.

This high value of θ suggests that agents are taking quite strong reaction to the lottery outcome. The high value of θ indicates two things. First, the room for taking the advantage of the aversion to regret does exist for the purpose of lottery promotion. This point was already well noticed in the literature. However, what has been neglected

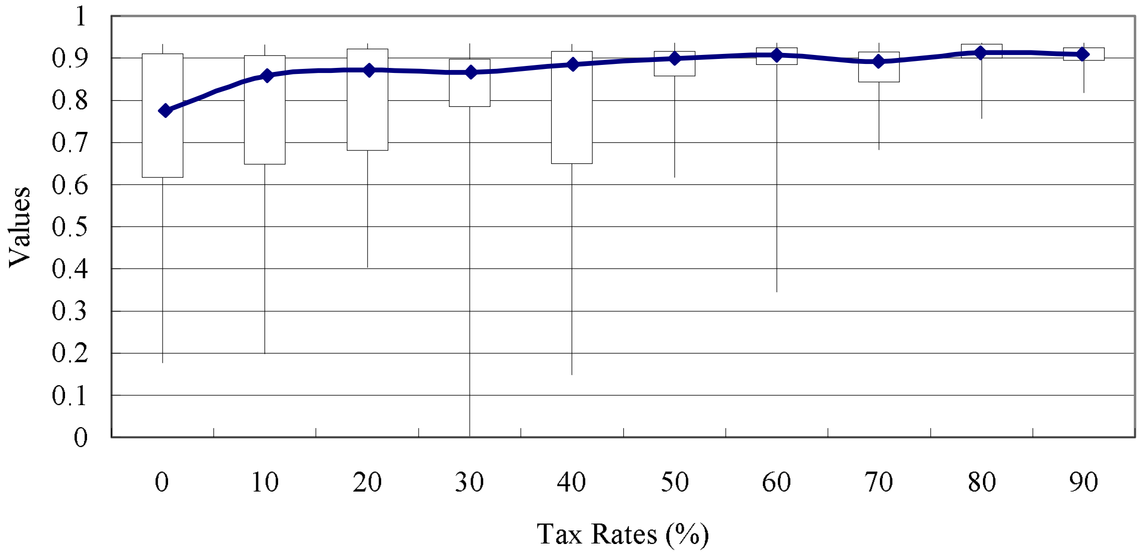

Figure 11: Value of the Regret Coefficient θ

is the dual role of θ: agents may also tend to decline to gamble with a *hidden wish* that no one will match the jackpot, and when that happens, they are just happier with their "prophesying" capability: "Yes! I just knew it." It would be interesting to see whether this trait emerges endogenously as a *psychological compensation*; after all, winning big prizes is difficult for most people. So, the *hidden wish* revealed by θ provides them a way to accept not being able to become a millionaire. This may partially explain why θ is significantly high for the cases where rollovers happen much more frequently, such as $\tau=$ 80%, 90% (see a very narrow box in Figure 11).[30]

Now, we can see why the removal of θ can enhance tax revenue, as founded in Section 5.1. Mainly, it is because that highly interdependent preferences has been developed, i.e. the parameter θ is significantly away from zero. For make the analysis easier, suppose $\theta = 1$. When there is no match for the top prize, non-performing agents will double their degree of happiness, from 200 to 400. This high degree of utility may even beat the winners of small prize, say the winner of the 4th prize. It is so because 35% of the prize is reserved for the 4th prize (Table 2), which is only 17.5% of sales if the tax rate is 50%. The winners of the fourth prize can only have 17.5 cents for per dollar purchasing. So, if 20% of the 5000 agents are gamblers, and each spends only one dollar for the ticket, then they together share the pool of 175 dollars. Given that, even though there is only one single winner of the fourth prize, there is still a short of 25 dollars to make this winner as happy as the non-performing agents.[31]

The analysis here shows the significance of evolving a positive θ: it nullifies the influence of those small winners. As a result, when the regret effect is removed (θ=0), the influential power of the small winners is brought back, and stimulates the participation to the market, which in turn levels up tax revenue.

[30] The findings that rollover frequency increases with the takeout rate will be shown later.

[31] The phenomenon that winners of small prize feel no better than the non-performing agents is not totally unrealistic. This can be confirmed from the fact many small-prize winners do not even bother to cash their reward. We, however, agree that the range of θ set here is a little exaggerating. But the purpose here is to make its effect easier to see.

6 Discussion: What does the GA learning mean?

Given the simulation results displayed above, it is high time to inquire a very fundamental question: *what does the GA learning mean?* This is a generic question shared by all kinds of agent-based simulation using the GA. To answer this question, we have to first notice that a possible optimal solution for all our agents in the lottery market is to take the *zero function* when the jackpot size is not high enough, i.e.,

$$\mu^* = \rho^*(J) = 0, \tag{15}$$

if J is not large enough. The solution is best in the sense that it maximizes the *risk-neutral expected utility* as specified in Equations (5) to (8).[32] The second thing to notice is that the fundamental work GA did in a social learning framework is simply to *propagate* those well-performed strategies based on the fitness function supplied by the user. If the fitness function is in line with the the utility function, then it is natural to ask whether the agents eventually find the optimal solution (15). In term of the discretized version of ρ, i.e., the participation vector $\vec{a}$, the optimal solution is

$$\vec{a}^* = (a_1^*, a_2^*, a_3^*, a_4^*) = (0, 0, 0, 0). \tag{16}$$

To distinguish this type of agents from other types, we shall call agents with solution (16) the *standard neo-classical agent.*[33] Our first question is then whether the solution (16) was propagated well enough to the entire market. This is a kind of the typical *convergence issue* addressed in may ACE simulations with the GA.

It is useful to look at the percentage of agents whose participation is in line with (15). Let N_t^* be the number of the standard neo-classical agents in the market, then the statistic $N_t^*/5000$ measures the density of neo-classical agents in the market. Since we have 25 simulations for each τ, each of which lasting for 500 issues, what drawn on Figure 12 is the box-and-whisker plot of $f_{500}^* = N_{500}^*/5000$. As before, dots of medians are connected in a line.

Figure 12 basically indicates the difficulty of propagating behavior (16). The percentage f_{500}^* is almost down to nil for most simulations when τ is less than 40%. While the further increase in τ does facilitate the propagation of the survival of the neo-classical agents, their influence is still confined to a rather limited extent. It is until the take-out gets to its maximum ($\tau = 80\%, 90\%$), they start to become a large group (one fifth to one third) of the survival agents. This result drives us to inquire what limits the survivability and propagation of the neo-classical agents. Or, put the question in the context of GA, why did the *presumably* well-performed strategy (16) fail to dominate?

The property that the behavior of the bounded-rational agents may not be able to converge to that of rational agents has been demonstrated in many ACE studies using GA. One key contributing factor to the divergence comes exactly from a life of bounded-rational agents, i.e., the *time-horizon* upon which agents react. In GA, this time-horizon is put into effect through the evaluation cycle. If the cycle is short, then the respective

---

[32] Walker and Young (2001) showed that when sales are large enough the expected returns of 1 dollar ticket comes to its maximum which simply equals $1 - \tau$. Therefore, even in this most favorable case, expected utility still goes down with the purchase of the tickets. It is true that Solution (15) may no longer be valid if different utility functions are applied. While this is indeed an interesting direction to look into further (e.g., Farrel and Hartley, 1998), this paper has not taken this line of arguments into account.

[33] In fact, what is needed is simply $a_i < \frac{1}{200} = 0.005$ for all $i = 1, 2, 3, 4$. This is so because in our simulation each ticket costs only 1 dollars, and agents' income is 200 dollars per period.

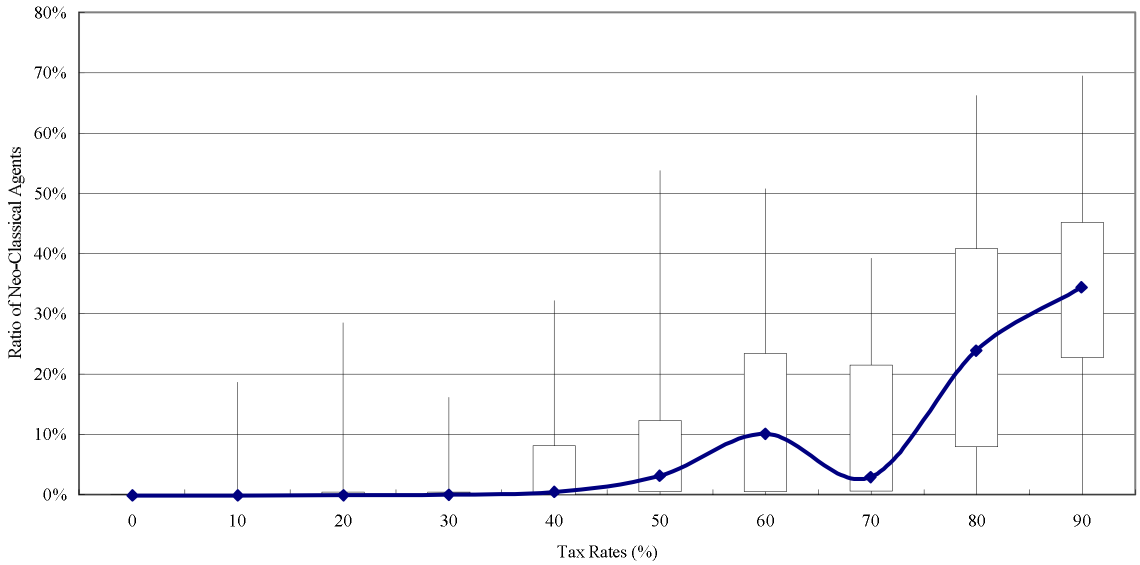

Figure 12: The Survival Rate of the Neo-Classical Agents

time-horizon is also short. The shorter the time-horizon, the more myopic the agents tend to be. Short time-horizons cause a problem well noticed by Lettau (1997), which is to be restated as follows. Agents in a setting of which the evaluation time-horizon is only *one period* are searching for

$$\vec{a}^{\dagger} = \arg\{\max_{a} \ U\}, \tag{17}$$

and Lettau (1997) has shown the non-equivalence between

$$E(\vec{a}^{\dagger}) = E(\arg\{\max_{a} \ U\}), \tag{18}$$

and

$$\vec{a}^{*} = \arg\{\max_{a} \ E(U)\}. \tag{19}$$

Lettau (1997) discussed the the two. That discussion basically applies to this paper.

For convenience, agents in the society can be decomposed into two groups: *gamblers* (performing agents) and *non-gamblers* (non-performing agents). Given the design of the lottery game, most gamblers will fail with a lower utility as opposed to the non-gamblers. Since the fundamental work which GA does is to *propagate* those well-performed strategies, the strategies used by these failed non-gamblers have no influence on shaping the forthcoming behavior. Nonetheless, a minority of gamblers, in particular those gamblers with aggressive participation, who are lucky to become the winners, gain utility which are significantly higher than those non-gamblers. These gamblers alone are persuasive enough to invite many followers to level up their participation. This explains why *gamblers can well propagate, even though most of them will fail.*

Non-gamblers can still exert some degree of influence to those losers, but their effect will be limited by the influence of the gigantic winners. But, if the gigantic winners do not show up (rollover), neo-classical agents will then have a better chance to fight back. Therefore, *the frequency of rollovers matters.* The more frequently the game rollovers, the more likely neo-classical agents can survive and propagate. To see this relation, Figure 13 depicts the percentage of rollover draws, or called the *rollover ratio*. Here, we see that

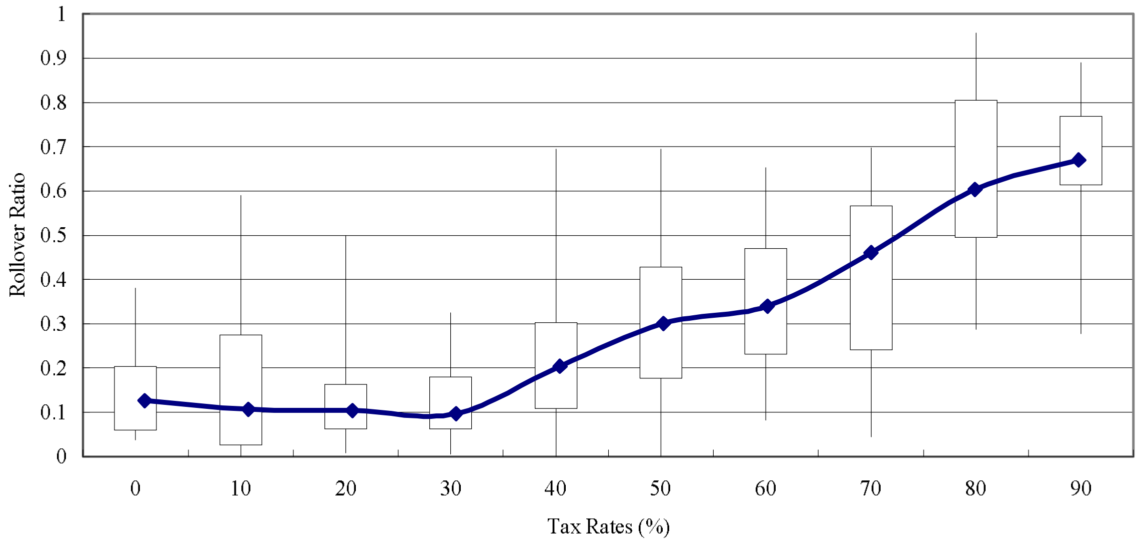

Figure 13: The Rollover Ratio

the rollover ratio roughly increases with the take-out rate. When τ comes near to the maximum, it is about 60% to 70% high, which means most of time the gigantic winners does not exist. As a result, neo-classical agents face much weaker survival pressure and can better propagate to a large proportion of market participants.

The explanation above also indicates an *asymmetric* account given between the minority of winning gamblers and the majority of the losing gamblers. The learning mechanism driven by the standard GA makes agents care very little of the losers, whatever strategies they used, even though they may use the same aggressive strategies of the winners. As a result, by not taking into account the losers' strategies, agents tend to give a *biased* or *over-optimistic* judgement of those aggressive gambling strategies. While the behavior observed is not rational in a strong sense, it is not at all atypical. Consider our experience with the so-called *recipes*. Sometimes, we were told and persuaded only when the recipe worked for some people, while how many times it failed for other people was largely neglected.

Here, we see the main distinction between the use of GA in engineering and the use of GA in social sciences. In most engineering applications, the fitness assigned to a specific strategy is deterministic, whereas in most ACE applications, it is non-deterministic. The problem of the non-deterministic fitness is usually tackled with by enlarging the evaluation time-horizon so that it is the average performance over many iterations by which a strategy is evaluated, as what was done in Lettau (1997).[34] However, the case of lottery presents an extreme situation: the reward for gamblers ranges from zero to gigantically high. Hence, a few more iterations will not help average out that windfall reward. As a result, *an essentially "hocus-pocus strategy", i.e., a strategy who work purely by luck, can still survive.*

The use of GA enables us to spell out the conditions upon which this property may sustain. First, human are *bounded rational* in the sense that they learn in a *biased* way

[34] Lettau showed that

$$\lim E(\vec{a}^{\dagger}) = \vec{a}^{*}. \tag{20}$$

by only drawing attention to the winners. Second, there is a *small probability* of having an *extremely high reward* of following that "hocus-pocus strategy" strategy. Notice that the probability can be extremely small. What matters here is that by *the law of large number* it will almost surely happens, at least for few people, if there are enough large number of followers. The case of lottery fit nicely to these two conditions. But, there maybe some other kinds of gambling or gambling-related behavior, e.g., crime, also fit these conditions. If so, the property implies these behavior will not die away and may self-form into a wave of propagation, even though rationally speaking it is not worthy of such doings.[35]

Now, come to the puzzle: why in our simulated data rollovers affect sales *negatively*, which contradicts to the most noticeable stylized fact of the lottery market. Again, this can be accounted by the way GA operated in this paper. Remember that our evaluation time-horizon is short to a single-period draw. Given the circumstances, this is what will happen. Suppose there was one and only one jackpot winner in the last issue (hence a regular draw for this issue). Let us trace his possible influential power in a framework of *tournament selection*. Since on each single draw we have a chance of 1/5000 to pick this jackpot winner, the chance of including at least one copy of the jackpot winner into a tournament is approximately 0.04 ($e^{200\times(1/5000)}$) if the tournament size is 200. There are 5000 tournaments (one for each individual), so on the average 200 individuals have the jackpot winner in their tournament and hence are under his influence. Since the jackpot winner tends to have a more aggressive participation $\vec{a}$. The aggressive strategy is, therefore, propagated to a large group of gamblers. This causes the rise of sales during the regular draw. On the other hand, the absence of the jackpot winner in the previous issue (rollover draw for the current issue) hampers the propagation of aggressive strategies. Instead, the conservative strategies which lead to low participation level dominates. Sales, therefore, fall in the rollover draw. Similarly, when the rollover draw extends, the jackpot prize accumulates. Therefore, a large jackpot prize, as a result of non-interrupting rollovers, also have adverse effect on sales. This explains the significantly negative t statistics and low $\bar{R}^2$s in Table 4.

The explanation above also suggests what may deviate the adaptive behavior in our artificial market from the adaptive behavior in the real market. First, it is the *speed of learning*. It is the extremely short time-horizon set in our GA leads to an apparently negative relation between rollovers and sales. If one extends the time-horizon and modifies the *fast learning* into a *slow learning* such that in every learning horizon regular draws is observed, then then immediate effect of no-match draw will be weakened, and it will facilitate the propagation of the aggressive strategies. So, if aggressive strategies tend to suggest to increase lottery participation with jackpot size, then the halo effect shall emerge again. This will be a direction for our next-stage simulations.

7 Conclusions

7.1 Remarks on the Findings

This paper introduces an agent-based computational model of lottery markets. In this model, agents' decision on lottery participation is not based on sophisticated calculation

[35] By this property, lottery participation or other similar gambling behavior can be addressed simply in a bounded-rational framework. Notice that here only standard utility function is used. There is no need to assume a *fun* from buying lottery tickets, neither is needed to assume a certain type of risk preference.

of winning odd but simply heuristics. The heuristics considered in this paper captures the two empirical phenomena known as the halo effect and the conscious selection of numbers. In addition, the empirical observation called the aversion to regret motivates an interdependent utility function of agents. The Sugeno style of fuzzy if-then rules are used to formalize agents' heuristics. Both the heuristics and preference are evolving over time via the canonical genetic algorithm.

We accept this simple ACE model as a starting point to conduct some initial evaluation of the impact of the lottery tax rate upon the the tax revenue. Two observations are made in this paper. First, the Laffer curve is observed, which suggests an optimal lottery tax rate, τ^*. Second, the τ^* can be sensitive to how agents are modeled. Simulations show that when the regret effect is moved away from agents' preference, the τ^* can go up. If so, the appearance of the interdependent utility function does have an implication on the design of the lottery game.

7.2 Direction for Further Studies

A number of interesting ideas on the lottery behavior can be studied following this line of research. Variations or extensions of this model can also be developed depending on the questions which we are pursuing. We propose a few directions to enrich our understanding of the lottery behavior.

First, it is the empirical relation existing between the *rollovers* and *sales*. Simple relation does not exists between rollovers and sales. While it is generally assumed that large size of rollovers will enhance the attractiveness of the lottery game, designing such a game may cause it more difficult to win, which in turn discourages the participation at the first place. As well put by Walker and Young (2001), "designing the game to maximise sales is a balancing act of making it hard enough to win to overcome the tedium but easy enough to win to avoid the intertemporal substitution. (p.702)" Alternatively speaking, the frequency and size of rollovers are *not* control variables (exogenous variables). Instead, they are endogenously generated simultaneously with sales by the design and other economic factors given in the vector $\mathcal{M}$. In the ACE model, both rollovers and sales are *entirely* endogenously determined. No assumption is needed on their stochastic behavior. We, therefore, can put the rollovers-sales relation in a more general framework, and examine how it is affected by different designs.

Second, a design issue which has not been fully explored in this ACE model is the distribution design $(s_0, s_1, ..., s_x)$. The simple fact is that practical lottery designs both large prizes as well as small prizes. The design $s_x = 1$ simply does not exist in the real world. So, to justify the distribution, it would be meaningful to simulate the our ACE model with different distribution. Nevertheless, this exploration can be done in a more meaningful way if we replace the *risk-neutral* preference with the *risk-averse* preference, because the former is only sensitive to *mean* of the prize distribution, whereas the latter is also sensitive to the change in the *high-order moments*, such as the *variance*, *skewness*, etc.

Walker and Young (2001) is probably the only paper which studies the effect of the distribution of the prizes on lottery sales. They found that the lottery ticket sales depends positively on the skewness of the prize distribution and negatively on the variance of the prize distribution. However, whether or not such existence of the distribution effect can lend support to risk averse preference is a difficult issue for Walker and Young to prove. Nevertheless, by simulating ACE models with both risk-neutral and risk-averse agents, one can identify whether the distribution effect has anything to do with gamblers' preference.

References

[1] Beenstock, M., E. Goldin, and Y. Haitovsky (1999), "What Jackpot? The Optimal Lottery Tax," The Hebrew University of Jerusalem, mimeo.

[2] Bell, D. (1982), "Regret in Decision Making under Uncertainty," *Operations Reserch*, **30**, 961–981.

[3] Bell, D. (1983), "Risk Premium for Decision Regret," *Management Science*, **29**, 1156–1166.

[4] Chen, S.-H. and C.-H. Yeh (2001), "Evolving Traders and the Business School with Genetic Programming: A New Architecture of the Agent-Based Artificial Stock Market," *Journal of Economic Dynamics and Control*, Vol. 25, pp. 363-393.

[5] Chen, S.-H. and C.-H. Yeh (2002), "On the Emergent Properties of Artificial Stock Markets," *Journal of Economic Behavior and Organization*, Vol. 49, pp. 217-129.

[6] Creigh-Tyte, S. and L. Farrell (1998), "The Economics of the National Lottery," *Working Papers*, Department of Economics, University of Durham (Web Site) (RePEc:dur:durham:190.).

[7] Farrel, L. and I. Walker (1996), "It Could Be You! But What's It Worth? The Welfare Gain From Lotto," *Working Paper*, Department of Economics, Keele University, Staffs ST5 5BG, UK.

[8] Farrell, L. and R. Hartley (1998), "Can Friedman-Savage Utility Functions Explain Gambling?" *Working Paper*, 98/02, Keele University.

[9] Farrel, L. and I. Walker (1999), "The Welfare Effects of Lotto: Evidence from the UK," *Journal of Public Economics*, **72**, 99–120.

[10] Farrell, L., E. Morgenroth, and I. Walker (1999), "A Time-Series Analysis of UK Lottery Sales: Long and Short Run Price Elasticities," *Oxford Bulletin of Economics and Statistics*, **61**, 513–526.

[11] Griffiths, M.D. and R.T.A. Wood (2002), "Lottery Gambling and Addiction: An Overview of European Research," *European Lotteries*, Psychology Division, Nottingham Trent University.

[12] Hartley, R. and G. Lanot (2000), "On the Design of Lottery Games," *Working Paper*, Department of Economics, Keele University.

[13] Hartley, R. and L. Farrell (2002), "Can Expected Utility Theory Explain Gambling?" *American Economic Review*, **92:3**, 613–624.

[14] Kahneman, D. and A. Tversky (1973), "Availability: A Heuristic for Judging Frequency and Probability," *Cognitive Psychology*, **5**, 207–233.

[15] Kahneman, D. and A. Tversky (1979), "Prospect Theory: an Analysis of decision under Risk," *Econometrica* **47**, 263–291.

[16] Loomes, G. and R. Sugden (1982), "Regret Theory: an Alternative Theory of Rational Choice under Uncertainty," *The Economic Journal*, **92**, 805–824.

[17] Mason, P.M., J.W. Steagall, and M.M. Fabritius (1997), "The Elasticity of Demand for Lotto Tickets and the Corresponding Welfare Effects," *Public Finance Review*, **25:5**, 474–490.

[18] McConkey, C.W. and W. E. Warren (1987), "Psychographic and Demographic Profiles of State Lottery Ticket Purchasers," *The Journal of Consumer Affairs*, **21**, 2, 314–327.

[19] Mikesell, J.L. (1994) "State Lottery Sales and Economic Activity," *National Tax Journal*, **47:1**, 165–172.

[20] Mikesell, J.L. and C.K. Zorn (1986) "State Lotteries as Fiscal Savior or Fiscal Fraud: A Look at the Evidence," *Public Administration Review*, **44**, 311–320.

[21] Mikesell, J.L. and C.K. Zorn (1987) "State Lottery Sales: Separating the Influence of Markets and Game Structure," *Growth and Change*, **18**, 10–19.

[22] Morgan, W.D. and J.D. Vasché (1979), "Horseracing Demand, Parimutuel Taxation and State Revenue Potential," *National Tax Journal*, **32**, 185–194.

[23] Morgan, W.D. and J.D. Vasché (1980), "State Revenue Potential of Parimutuel Taxation: A Comment," *National Tax Journal*, **33**, 509–510.

[24] Morgan, W.D. and J.D. Vasché (1982), "A Note on the Elasticity of Demand for Wagering," *Applied Economics*, **14**, 469–474.

[25] Paton, D., D.S. Siegel, and L. Vaughan (2002), "The Impact of Taxation on the Demand for Gambling," *Economic Division Working Paper Series*, N.05–2002, Nottingham University Business School.

[26] Purfield, C. and P. Waldron (1999), "Gambling on Lotto Numbers: Testing for Substitutability or Complementarity Using Semi-Weekly Turnover Data," *Oxford Bulletin of Economics and Statistics*, **61**, 4, 527–544.

[27] Scoggins, J.F. (1995), "The Lotto and Expected Net Revenue," *National Tax Journal*, **5**, 61–70.

[Statman(2002)] Statman, M. (2002), "Lottery Players/Stock Traders," *Financial Analysis Journal*, January/February.

[28] Walker I. (1998), "The Economic Analysis of Lotteries," *Economic Policy*, **13**, 27, 357–402.

[29] Walker, I. and J. Young (2001), "An Economist's Guide to Lottery Design," *The Economic Journal*, **III** (Nov.), 700–722.

[30] Von Neumann, J. and Morgenstern, O. (1947), "Theory of Games and Economic Behavior," Ed.2. Princeton University Press, Princeton, N.J.