

Nicolás Garrido and Luigi Mittone

**A Description of Experimental Tax Evasion
Behavior Using Finite Automata: the case of
Chile and Italy**

CEEL Working Paper 9-08

Computable and Experimental Economics
Laboratory

Via Inama, 5 38100 Trento, Italy

<http://www-ceel.economia.unitn.it>
tel. +39.461.882246

A Description of Experimental Tax Evasion Behavior Using Finite Automata: the case of Chile and Italy

Nicolás Garrido* Luigi Mittone[†]

October 9, 2008

Abstract

In this paper, we use a Moore Automata with Binary Stochastic Output Function in order to explore the extensive decision regarding tax evasion made by subjects in experiments run in Chile and Italy. Firstly, we show how an hypothesis about subject behavior is converted into an automaton, and how we compute the probabilities of evading for every state of an automaton. We use this procedure in order to look for the automaton which is able to anticipate the highest number of decisions made by the subjects during the experiments. Finally, we show that automata with few states perform better than automata with many states, and that the bomb-crater effect described in [1] is a well identified pattern of behavior in a subset of subjects.

1 Introduction

”As economists we are confronted with a choice between waiting for a satisfactory description of the procedure of human decision making and analyzing somewhat artificial models capturing certain elements of ‘bounded rationality’” Ariel Rubinstein [2] [p.83]

In this paper we use the theory of finite automata starting from a set of behavioral data collected in two set of experiments on tax evasion run in Italy and in Chile.

The neoclassical theory of tax evasion [3] [4] is based on the assumption of a traditional full rational agent who decides her level of tax compliance accordingly with a standard process of expected utility maximization. This theoretical approach has been widely criticized mainly moving from an experimental empirical ground. The extraordinary wide latitude of the experimental economics literature in this field makes impossible to summarize here the richness of the empirical findings (for a recent survey see Eric Kirchler [5]).

In extreme synthesis one could say that the standard theory fails to account in predicting the agents’ reactions induced by the standard parameters of the decisional problem - a typical example of this kind is the weakness of the theory in explaining the tax payers reactions to an increase in the tax rate -. Moreover the standard theory is inadequate in forecasting the tax payers’ response to the

*Universidad Católica del Norte. Nuclei of the Millennium Science Initiative Program ”Regional Science and Public Policy”, Chile, ngarrido@ucn.cl

[†]Università di Trento. Dipartimento di Economia, Computable and Experimental Economic Laboratory (CEEL)

inclusion of psychological factors in the task problem - e.g. including a social blame effect or other forms of psychological constraint -.

Some quite recent attempts to overcome the limits of the standard approach to tax evasion have been carried out by substituting the Expected Utility Maximization Theory (EUMT) with the Prospect Theory (PT) by Kahneman and Tversky [6]. In spite of its behavioral foundations the Prospect Theory still belongs to the class of theories based on a perfectly rational agent. The only apparent concession to bounded rationality in the theoretical machinery of PT is limited to the allowance of some inconsistency in the risk attitudes. The PT agent re-models her risk attitude accordingly with a reference point which works as a psychological dividing wall between risk aversion and risk propensity.

The psychological-behavioral foundations of PT can help in explaining the tax payers' inconstancy reported by many repeated choices experiments but is still unable to forecast in a satisfactory way the full complexity of the factors that can affect the decisional process. This is clearly a paradox: from one side both the two most important among the general parsimonious theories applied to tax evasion fail to take into account the complexity of the decisional process and from the other side these theories are grounded on a substantially perfectly rational agent who does not exist in the real world.

The attempt carried out in this paper is to go into the opposite direction. We assume an agent who is strongly rationally bounded, even more limited than the real agents are, but at the same time we use a quite sophisticate analytical approach to model her behavior.

Using finite automata for explaining human decision making, is equivalent to say that participants' decisions depend on their current conditions or state. An individual has different states and she changes from one state to another according to external events. This idea closely resembles the original Simonian definition of bounded rational agent. To decide accordingly to a bounded "local" set of information implies the discard of the standard view of a perfectly rational optimizing homo economicus. If an individual has evaded but she was not audited she might be less willing to evade the next time than a subject who has been audited and punished. It is worth noticing that this shifting of the agent's attitude towards evasion due to past experiences is not compatible with a standard view of perfect rationality. What has happened in the previous round should not affect the tax payer decision process because if she would be perfectly rational then she has had already computed all her future choices. More precisely the perfectly rational tax payer chooses on the basis of a general probabilistic inter-temporal plan built on a complete preference mapping of all the possible alternative probabilistic weighted outcomes. Contrary to this picture, in our setting, we assume that the probability of evading depends on the current subjective state of the tax payer who is determined by the local status quo.

The idea that finite automata theory may be useful for modeling human making decision is not new. The just quoted Rubinstein originally developed a research agenda where automaton was used for modeling bounded rationality and Romera [7] uses finite automata to represent mental models.

Another similar view of a locally determined decision maker which is closer to our theoretical design is suggested by Selten [8]. In his Aspiration Adaptation Theory the Seltenian agent chooses her actions by considering a limited set of decisional dimensions (the aspiration level) described by discrete scales of measure which are determined by the specific local state of nature. The same agent can choose opposite actions depending from the local bond of a dynamically adaptive network of aspirations described in a discrete finite set of alternatives. Moving from one bond to another means changing from an aspiration level to another without having a complete picture of the whole set of aspirations. In this sense we can imagine a possible (extreme) adaptation of the Aspiration Adaptation Theory to our approach. In our setting the relevant dimensions of choice in the tax payer problem is represented by the two states “has been audited”, “has not been audited” and the tax payer decides only accordingly to the characteristic of her current state.

It is worth noticing that the Aspiration Adaptation Theory applies very well as a general descriptive tool for modeling the tax payer behavior and can also been applied to the experimental design used here. An aspiration level in Selten’s definition can be described as a vector $a = (a_1, \dots, a_m)$ where a_j is the partial aspiration level for a generic G_j goal. A goal for a tax payer could be the amount of income “saved” from taxes or the psychological cost implied by the decision to evade due to social blame. In the theoretical setting of the Adaptation Aspiration Theory the tax payer should move from an aspiration level to another accordingly with the starting point and using a sort of backward looking logic that pushes her from one local aspiration level to another one. The choice of the new aspiration level depends on the local informational bundle which in our example could be represented by the amount of disposable income, the value attributed to the moral constraints, etc. An exhaustive description of an aspiration level applied to tax evasion goes beyond our aims so what we shall take from the Adaptation Aspiration Theory is more a methodological hint than a complete theoretical frame.

Our model does not explain the full process of rationing made by the agent before of making the decision. Instead of this, we characterize the states where the subjects are more willing to evade or not. The willingness to evade will be represented as a probability.

We use a type of finite automaton named Moore machine. This machine consist of a finite set of states, one of them includes an initial state, an output function and a transition function. We partially modify the output function; instead of depending only on the current state, the decision of evading or not depends on the current state and its related probability of evasion.

As the number of states of the machine increases, we obtain a better explanation of the subjects behavior. However, in this case we obtain automata with a high number of states. These type of automata do not produce useful theory for explaining subjects decisions. Therefore, we want to model the behavior of the agents with the lowest level of complexity. Having a simple automaton allows us to

have good explanation of subject decisions. Although several measures of complexity for the automata have been suggested in the literature (See [2]), we will use a fairly naive definition of complexity: what counts is the number of states in the machine.

Our main conclusion is that taking into account the extensive decision on evasion made by the subjects, we distinguish two patterns of behaviors: on the one hand there is a group of individuals who behave honestly, paying all their taxes all the time. On the other hand we confirm the relevance of the so called "bomb crater effect" [1] which predicts that individuals are more willing to evade right after being audited (see section 3).

In the next section we will formalize the Moore Automata with Binary Stochastic Output Function for capturing the subjects behaviors. After the formalization we show two methods for estimating the probabilities in every state: first we use different probabilities for discriminating the subjects behavior during the experiments run in Chile and Italy. Second, we optimize the probabilities for explaining only with one automaton the behavior of all the subjects in a given experiment. After this, we conclude describing the results and future work to be done.

2 The Moore Automata with Binary Stochastic Output Function

In the theory of computation, a Moore machine is a finite state automaton where the outputs are determined only by the current state and does not depend directly on the input. The standard state diagram for a Moore machine includes a deterministic output signal for each state.

We introduce a variation in the output function of the Moore automata: instead of producing a deterministic output signal, the output function can produce either of two values 0 or 1. Every state s has a probability p_s of producing a 1 and a probability $(1 - p_s)$ of producing a 0.

Thus, the Moore machine with a binary stochastic output function can be defined as a 7-tuple $\Gamma = \{S, S_0, P, \Sigma, \Lambda, T, G\}$ consisting of the following objects:

- a finite set of states (S),
- a start state (also called initial state) S_0 which is an element of (S),
- a set of probability values (P). Every state $s \in S$ has a probability $p_s \in P$. The initial state has a probability $p_{S_0} \in P$. The numerosity of the set P is equal to the numerosity of S plus 1, i.e. the initial state,
- a finite set called the input alphabet (Σ),
- a finite set called the output alphabet ($\Lambda = \{0, 1\}$),

- a transition function ($T : S \times \Sigma \rightarrow S$) mapping a state and an input to the next state
- an output function ($G : S \rightarrow \Lambda$) mapping each state and its probability in (P) to the output alphabet as follows

$$G(p_s) = \begin{cases} 1, & \text{if } \epsilon < p_s; \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

where ϵ is a uniform random number between 0 and 1.

This automaton produces a binary string of 0 and 1 according to the visited state and the probability of that state. Each 1 and 0 is the prediction made by the automaton about the subject decision for the next period. If the probabilities of all the states is 0.5, the automaton will produce a random string of 1's and 0's.

The design chosen for our automata is quite close to the Selten's idea of dynamically adaptive aspiration plans. Our automata agent can be described as a decision maker who chooses her actions accordingly with a given aspiration level. In our model the local dependence of the aspiration level is mimicked by the probability state. In Selten's Theory probability has no role because the Seltenian decision maker is fully deterministic but our concept of local probability must be interpreted here as a way to define the subjective dimension of the aspiration plans. More precisely the transition function, combined with the output function (which incorporates the state probability), describes the decision making process which drives our automata from one bond to another of the decisional network. It is worth noticing that the states of our automata can be interpreted as aspiration levels.

3 Data from the Experiments: extensive and intensive decisions

The experiments discussed in this work have been carried out using an identical experimental design - software, instructions, payoffs structure - in Italy at the CEEL laboratory of the Trento University and in Chile at Northern Catholic University. The experimental design has been originally developed in a previous work by [1] and the Italian data here reported have been extracted from the same paper. The baseline experiment is based on a repeated choices setting. Two alternative treatments of the baseline experiment are here reported as well. Treatment one introduces a mechanism of redistribution of the tax yield among the participants while treatment two uses the tax yield collected to finance the production of a public good which is consumed outside the strict group of the participants. The hypothesis tested with treatment one and treatment two is that sentiments of other regarding can contrast tax evasion *ceteris paribus*. The baseline session as well as the treatments sessions have been run using a computer-aided game. Thirty undergraduate students participated in each session, 15 men and 15 women. All the experiments were of the same length (60 rounds). The parameters entered in

the experiment are the followings:

- income - 0,51 Euro cents from round 1 until round 48, then 0.36 Euro cents;
- tax rate - 20 from round 1 until round 10, then 30 from round 11 until round 30, and finally 40 from round 31 until the end;
- tax audit probability - 6 from round 1 until round 21, then 10 from round 22 until round 40, and finally 15 from round 41 until the end;
- fees - the amount of the tax evaded plus a fee equal to the tax evaded multiplied by 4.5; the tax audit had effect over the current round and the previous three rounds.

To approximate a real life situation more closely, the tax audit was extended over a period of four rounds in the base experiment. For this reason, and as the lottery structure changed during the experiment, computation of the expected value from evasion was rather complex. The lottery structure for the experiment was always unfair which means that a risk neutral tax payer should always pay the entire amount of tax due. (see [1] for the details) The experimental subjects have recruited either in Italy and in Chile through announcements on bulletin boards. During the experiment the players were kept separated so that they cannot communicate in any way. The relevant pieces of information have been communicated only via the computer screen, which showed the following items:

1. the total net income earned by the player since the beginning of the game,
2. gross income in the active round,
3. the amount of taxes to pay in the active round,
4. the number of the active round.

The subjects were divided into two groups, and they underwent a fiscal audit in the same rounds (specifically rounds 13, 31, 34, 48, 54, 58 for the first group, and rounds 3, 24, 27, 40, 46, 50 for the second group). The two sequences have been randomly extracted accordingly with the rules of the design (i.e. using the probability values corresponding to each of the three sub periods of the game). The changes in the audit probability have been communicated to the participants with a warning on the computer screen which informed them that the audit probability would change after three rounds. Each participant every round must write, using the computer keyboard, the amount of money that she has decided to pay and then she must wait for being informed on the result of the audit extraction. In the following section a brief overview of the data obtained from both the Italian and Chilean experiments is reported. The main result is that the behavioral patterns into the two countries are very similar and in particular the so called "bomb crater effect" [1] Kirchler et al. appears as a robust phenomenon in all the samples.

For the sake of making easier the reference below, we denote the set of experiments for Chile and Italy by $\Xi = \{e_{c1}, e_{c2}, \dots, e_{c6}, e_{i1}, e_{i2}, \dots, e_{i6}\}$. Experiments e_{c1} y e_{i1} are two experiments with the same experimental design, but they differ because the former was run with chileans subjects whereas the second with italians one. Thus the experiments are

e_{c1}, e_{i1} : baseline experiment with group 1 for Chile and Italy.

e_{c2}, e_{i2} : baseline experiment with group 2 for Chile and Italy.

e_{c3}, e_{i3} : mechanism of redistribution of the tax yield with group 1 for Chile and Italy.

e_{c4}, e_{i4} : mechanism of redistribution of the tax yield with group 2 for Chile and Italy.

e_{c5}, e_{i5} : tax yield to finance the production of a public good with group 1 for Chile and Italy.

e_{c6}, e_{i6} : tax yield to finance the production of a public good with group 2 for Chile and Italy.

In every experiment e . are 15 subjects. Each of the 180 subjects in Ξ made decision on the amount of payment during 60 periods. Thus, any agent j has produced a history of 60 decisions. We are interested only on the extensive decision of evading. If evasion is represented with a 1, and not evasion with a 0, the decisions made by an agent j in an experiment, can be represented by a binary string h_j of size 60.

3.1 The case of Chile

Looking at the extensive decision made by the subjects, in Figures 1, 2, 3 and 4 we plot the number of subjects deciding to evade in every period for the experiments made in Chile and Italy. In the horizontal axis there are the periods of the experiment, and in the vertical axis the number of subjects who decide to evade. The vertical bars represent the periods where all the subjects were audited. Notice that the difference between the group 1 and 2 is on the distribution of controls. In the group 2, the sequence of controls is shifted toward the beginning of the experiments.

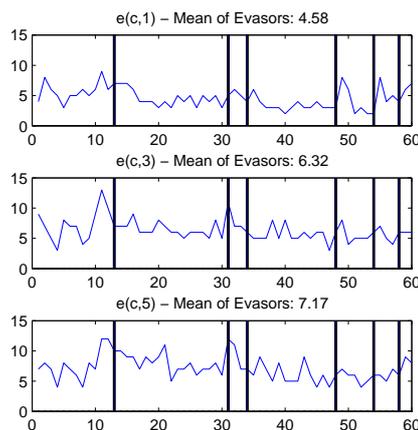


Figure 1: Extensive Decisions for Group 1 in Chile

There are some characteristics which are important to point out for the cases of Chile. In the first experiment, with the first group in Figure 1 there is an greater number of subjects evading after the government control. This seems to confirm the bomb crater effect. However, in the other two experiments for the first group, the effect is unclear: sometimes there are less evaders and occasionally more evaders after a control.

For the second group, in Figure 2 it is difficult to anticipate a clear aggregate patter of behavior. Looking at the average number of evaders, there is more evasion in the second group than in the first group.

In all the cases, there is no event that coordinate all the subjects for evading or not. In the population of 15 subjects, the changes on the decision of evading or not from one period to the other was never greater than 5. Moreover, in every period in all the experiments there were at least 3 subjects evading.

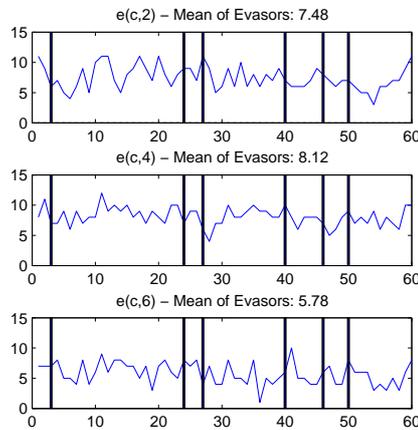


Figure 2: Extensive Decisions for Group 2 in Chile

3.2 The case of Italy

Looking at the number of evaders in every period for the experiments in Italy, the first group shows more irregular patterns than the chilenean case. In another words, it seems to be that italians sometimes coordinates to make the same decision.

In experiment 1 and 3 there is an stable number of evaders at the beginning. However the redistribution experiment, produces high variability of evaders at the beginning.

The average number of evaders in all the experiments is lower for the second experiment. The induced effect of the public good has produced a reduction of evaders in the italian case.

For the case of Italy and Chile we can observe that on the aggregate data, it is difficult to observe the existence of the bomb crater effect. In what follow we will show that in aggregate this effect is

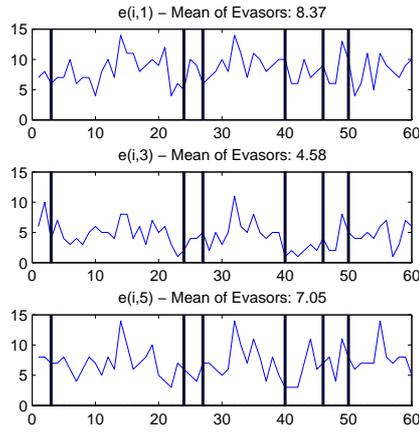


Figure 3: Extensive Decisions for Group 1 in Italy

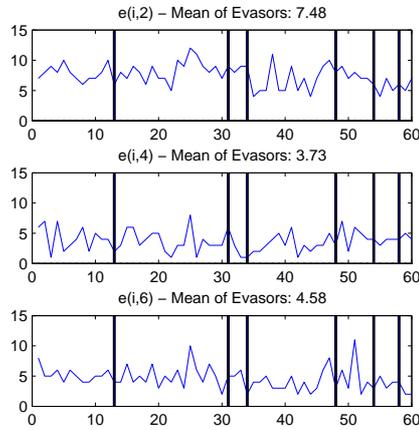


Figure 4: Extensive Decisions for Group 2 in Italy

indistinguishable because not all the subjects behave according to that effect.

4 Using automata for characterizing subject behaviors

In this section we search for an automaton or a set of automata for making the best prediction of the decision of evading of the subjects in the experiments.

We distinguish two possible strategies for accomplishing such a goal: first a computationally intensive strategy, and second the strategy of searching for the minimum set of automata explaining a high proportion of subject behaviors.

The computationally intensive strategy is to build an automaton for each subject in every experiment¹. Each of these automata could have sixty states, one for each period of time.

¹We have data from two countries. In every country we ran three experiments, and in every experiment there are two groups with fifteen subject each. This would mean that our theory had $2 * 3 * 2 * 15 = 180$ automata.

Although the computational intensive strategy provides good prediction on subject behaviors, it does not give us any useful abstraction for a theory on fiscal evasion behavior. Epistemological and ontological parsimony is a basic requirement to make such a contribution. The construction of 180 automata does not provide good insight to pursue this goal.

The second strategy, searching for a set of minimum automata for explaining the highest proportion of subject behavior, is more challenging and eventually give us the chance of making theory about subject behavior on tax evasion.

In this strategy we search for a set of automata, (ideally only one automaton) with the highest success rate at predicting the extensive behavior of all the subjects in a given experiment.

4.1 Computing the success rate of a set of automata

This procedure explains how do we obtain the number of states of an automaton, how do we compute the probability of evading in every state and finally how do we compute the success ratio of the automaton.

1. We start with an hypothesis H about the behavior of the subjects in a given experiment $e. \in \Xi$,
2. we transform the hypothesis into a finite automaton Γ_H . The hypothesis determines the number of states n of the automaton,
3. for any single subject j we compute a vector of visited states V_j . The visited state vector V_j is a $(n \times 1)$ vector where every position $V_j(i)$ for $i = 1..n$ is the number of times that the agent j visited the state i in his history h_j .
4. for every subject j we compute the vector of evasions E_j . This vector is a $(n \times 1)$ vector where the position $E_j(i)$ for $i = 1..n$ count how many times the subject j evaded after of being in the state i in h_j .

Thus, if an hypothesis produces an automaton with 8 states, and a given agent j has $V_j(3) = 5$ and $E_j(3) = 0$ it means that the agent j was in the state 3 five times and she never evaded in that state.

5. we compute the vector of probability Π_j . This is a vector of $(n \times 1)$ dimension defined as

$$\Pi_j(i) = \begin{cases} \frac{E_j(i)}{V_j(i)}, & \text{if } V_j(i) > 0; \\ 0, & \text{otherwise.} \end{cases} \quad \text{for } i = 1..n \quad (2)$$

Thus for every agent j we have a vector Π_j showing what is the empirical probability of evading for the agent j given that he is in an state i with $0 < i \leq n$.

6. we use the vector Π_j of all the agents in a given experiment to create a set of K clusters. Each one of this cluster represents a group of subjects with similar vector Π of probabilities of evading in each states. The clusters create a partition of subjects $\Omega = \{\omega_1, \dots, \omega_K\}$. Thus for every subject $j \in e$. he can be classified in one an only one of the subsets ω_k with $1 \leq k \leq K$
 7. we construct a set of automata $\Gamma(e.) = \{\Gamma_1, \Gamma_2, \dots, \Gamma_K\}$. All the automata in Γ have the same number of states. They only differ in their probability vector P^2 . The vector of probability of the automaton Γ_k is computed as $P_k(i) = \frac{\sum_{j \in \omega_k} \Pi_j(i)}{\#\omega_k}$ for $i = 1..n$. The function $\#\omega_k$ returns the number of subjects classified in the cluster ω_k .
- The set of automata $\Gamma(e.) = \{\Gamma_1, \Gamma_2, \dots, \Gamma_K\}$, represents the behavior of all the subjects in an experiment $e.$. Every automaton Γ_k represents a set of subjects with the same behavior.
8. For every subject $j \in e$. we select his history of extensive decision h_j and the automaton which represent him Γ_k where k is the automaton that capture the behavior of j . Formally k is such that $j \in \omega_k$.

We simulate during 60 periods the automaton Γ_k for anticipating the decision made by the agent j along her history h_j . This generates the prediction \hat{h}_j . The prediction is the result obtained from the output function $G \in \Gamma_k$. The success rate is the proportion of correct prediction made by the set of automata $\Gamma(e.)$ out of all the possible evasion decisions made by all the subjects in $e.$.

More precisely the success rate predicting the evasions of agent j as χ_j is

$$\chi_j = 1 - \frac{\sum_{t=1}^{60} |\hat{h}_j(t) - h_j(t)|}{60} \quad (3)$$

where $|\cdot|$ is the absolute value function. The average value across all the subjects represented by an automaton is its success ratio.

In the next section we will develop an example of our method for computing the success rate for a given set Γ_e .

4.2 The basic hypothesis: a three state automaton

1. Suppose the following hypothesis H_0 :*The decision of evading depends on whether the subject was audited or not in the previous period.*
2. The automaton for representing this hypothesis is $\Gamma_{H_0} = \{S, S_0, P, \Sigma, \Lambda, T, G\}$ and Figure 5 represents it.

²There is a difference between the probabilities Π y P . The probabilities Π are the empirical probabilities obtained in the previous steps, whereas P is the vector of probability related to the output function of every automaton in $\Gamma(e.)$.

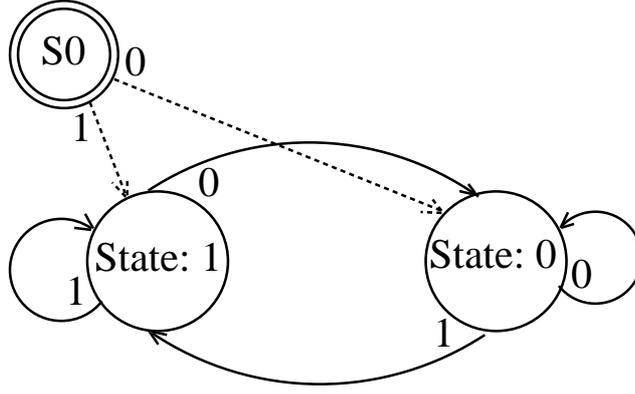


Figure 5: The three state automaton

The automaton has two states $S = \{0, 1\}$ and a start state S_0 . Thus, in order of representing the hypothesis we require only three states: the 0 in S represents the state of not audited in the previous period. The 1 in S represents the state of audited in the previous period.

The input alphabet of this automata is $\Sigma = \{0, 1\}$, where the 0 means not audited and the 1 means audited.

The output alphabet is $\Lambda = \{0, 1\}$, where 0 means it will not evade and 1 means that in the next move the subject will evade.

The transition function for all the states is defined as follows $T(S_0, 0) = 0; T(S_0, 1) = 1; T(0, 0) = 0; T(0, 1) = 1; T(1, 0) = 0, T(1, 1) = 1$. The first parameter of the transition function represents the current state of the automaton. The second parameter represents what was read from the input. For instance, $T(1, 0) = 0$ means that the automaton is in state 1, i.e. audited in the previous period, and it reads the symbol 0, i.e. the subject was not audited in the current period, the automaton change to the state 0, i.e. not audited in the previous period.

The output function is probabilistic and depends on the probability vector P . The values of the probability vector P will be computed below.

3. we compute the vector of visited state V for the 15 subjects in the experiment e_{c1} . For example, the visited state vector of subject 1 in e_{c1} is $V_1 = \{54, 6, 1\}$. Indeed, during the experiments there are 6 audits. This is why subject 1 visited the state: *not audited* 54 times, the state *audited* 6 times and finally the initial state S_0 one time.
4. we compute the vector of evasion E for the 15 subjects in the same experiment e_{c1} . For the first subject, the vector is $E_1 = \{16, 1, 0\}$. This means that the subject 1 evaded 16 times when he was in the state *not audited*, he evaded only one time when he was in the state *audited* and finally the subject did not evade in the initial state.

Table 1: Probability vectors of the automata Γ_1 and Γ_2 for e_{c1}

| State | P_1 | P_2 |
|-----------------------|-------|-------|
| Initial State S_0 | 1 | 0.4 |
| Audited State "1" | 1 | 0.44 |
| Not Audited State "0" | 1 | 0.25 |

5. we compute for all the subjects their probability vector Π_j for $j = 1..15$. For instance, the probability vector for subject 1 is $\Pi_1 = \{0, 2963; 0, 1667; 0\}$.
6. using all the vectors Π_j for $j = 1..15$, and deciding $K = 2$, we divide using cluster techniques the group of 15 subjects in two³. For instance in this case, the optimal partition of the subjects in e_{c1} is given by $\Omega = \{\omega_1 = \{3\}, \omega_2 = \{1, 2, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15\}\}$
7. now we build the two automata, Γ_1 and Γ_2 and their probabilities vectors are given by $P_1 = \{1, 1, 1\}$ and $P_2 = \{0.2564, 0, 4487, 0, 52\}$.

The automaton Γ_1 represents the individual in the first cluster. The subject 3 has evaded in all her decisions. The automaton Γ_2 represents all the other individuals. According to its probability vector P_2 the subjects evade more when they were in the state of *audited*.

8. using both automata we generate the predictions \hat{h}_j for $j=1..15$. Notice that for the subject 3 we simulate her predictions using Γ_1 whereas for all the other subjects we use Γ_2 . Taking into account that the simulations of the predictions have an stochastic component in the output function, we run Monte Carlo experiment in this step.

The success ratio of the hypothesis H_0 using two automata $\Gamma = \{\Gamma_1, \Gamma_2\}$ is 0.6328.

The probability P of every automata provide information about the behavior of the subjects represented by the set of automata Γ_{H_0} .

In the step number 7 we show the probabilities of evading by the two automata. The first automaton represent a subject who evade all the time. This subject just kept evading independently of the state where he was found. On the other hand, Γ_2 represent the average behavior of the subjects in the experiment e_{c1} . Notice that the probability of evading after being audited is almost twice the probability of evading if the subject was not audited. This is the bomb crater effect already mentioned.

In Figure 6 we present a complete exploration of the performance of the hypothesis H_0 which is represented with a three state automaton. The graph on the left has in the horizontal axis the number of clusters or automata with different probabilities, and in the vertical axis there is the success ratio of the automata. We explore the performance from 1 to 15 automata in each experiment, and for all the countries. Notice that more than five automata with different probabilities the improvement in the

³The decision of the number of clusters is arbitrary

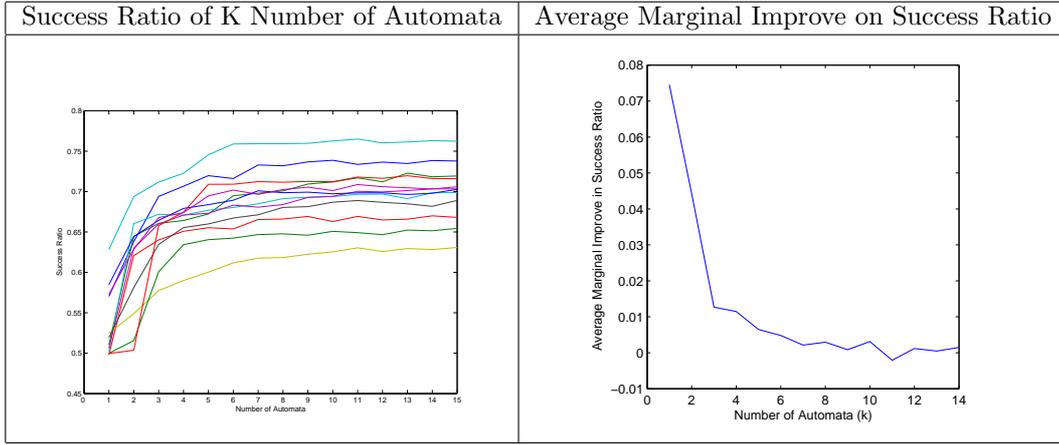


Figure 6: Three State Automata with Different Number of Automata

Table 2: Probability vectors of the automata $\Gamma_{H_0}(e_{i5}) = \{\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5\}$

| State | P_1 | P_2 | P_3 | P_4 | P_5 |
|-----------------------|----------|---------|----------|----------|----------|
| Initial State S_0 | 0 | 0.32 | 1 | 0 | 0.6 |
| Audited State "1" | 0.5 | 0.75 | 0.33 | 0.166 | 0 |
| Not Audited State "0" | 0.74 | 0.30 | 0.13 | 0.64 | 0 |
| Representation | 1 (0.06) | 6 (0.4) | 2 (0.13) | 1 (0.06) | 5 (0.33) |

success ratio is negligible. Indeed the graph on the right shows that after five automata the marginal improvement on the average success ratio is negligible.

The experiment e_{i5} is the experiment which obtain the highest success rate. In order to obtain information from its performance we analyze its behavior with more detail. We run the analysis with $K = 5$, or with five automata with different probabilities which produces a success rate of 0.7455.

Table 2 shows the probabilities of each automaton together with its degree of representativeness, i.e. how many subjects the automaton represents. Automata Γ_2 and Γ_5 have the highest degree of representativeness. The former represents six subjects and the later five.

The group represented by Γ_2 in the third column, shows the effect of the bomb crater effect and the group represented by Γ_5 are subject who no matter the state where they are, they pay all their taxes.

The two subjects represented by Γ_3 evade few times, however they behave according to the prediction of the bomb crater effect.

4.3 Exploring hypothesis

Following once more to the brief theoretical introduction and looking to the data collected from the experiments one can conclude that it might be the case that there is not such a thing as the best hypothesis for explaining how subjects behave in every experiment.

Table 3: Hypothesis that explain better the behavior of each experiments with K automata

| Experiment | Hypothesis | Success Ratio | (K) Clusters |
|-------------------------------|------------|---------------|--------------|
| e_{c1} Base Line, Group 1 | H_0 | 0.6931 | 3 |
| e_{c2} Base Line, Group 2 | H_2 | 0.6644 | 3 |
| e_{c3} Redistribution, g. 1 | H_0 | 0.6614 | 3 |
| e_{c4} Redistribution, g. 2 | H_3 | 0.6660 | 3 |
| e_{c5} Public Good, Group 1 | H_0 | 0.6455 | 3 |
| e_{c6} Public Good, Group 2 | H_0 | 0.5752 | 3 |
| e_{i1} Base Line, Group 1 | H_2 | 0.6557 | 3 |
| e_{i2} Base Line, Group 2 | H_0 | 0.6686 | 3 |
| e_{i3} Redistribution, g. 1 | H_0 | 0.6634 | 3 |
| e_{i4} Redistribution, g. 2 | H_3 | 0.7549 | 5 |
| e_{i5} Public Good, Group 1 | H_0 | 0.601 | 3 |
| e_{i6} Public Good, Group 2 | H_0 | 0.6643 | 3 |

Using the available information, we propose a set of hypothesis and we explore how do they performance at explaining the subjects behavior. In order to do that, we compute the success ratio obtained for each hypothesis. The set of hypothesis that we use are,

H_1 : The decision of whether to evade or not depends on whether the agent was audited or not during the previous period and if when he was audited he was caught evading.

H_2 : The decision of whether to evade or not depends on the whether the agent was audited or not during the last two periods.

H_3 : The decision of whether to evade or not depends on whether the agent was audited or not during the last two period and if when he was audited he was caught evading.

H_4 : The decision of whether to evade or not depends on the whether the agent was audited or not during the last three periods.

H_5 : The decision of whether to evade or not depends on the whether the agent was audited or not during the last fourth periods.

H_6 : The decision of whether to evade or not depends on the whether the agent was audited or not during the last three periods and if when he was audited he was caught evading.

In Table 3 we present the experiments, the hypothesis that explain them better and the number of clusters that we use for obtaining the success ratio.

The success of the hypothesis H_0 suggest us that longer histories, in the sense of taking into account the past of decisions, does not improve the success ratio of the automata.

Each automata representing the hypothesis in Table 3 characterize the behavior of the subjects in the experiments. For the sake of space we explore with more detail experiment e_{i4} , because we

Table 4: Automata Γ_{H_3} using K=5 clusters

| t-2 | t-1 | x | Γ_1 | Γ_2 | Γ_3 | Γ_4 | Γ_5 |
|-----|-----|---|------------|------------|------------|------------|------------|
| 0 | 0 | 0 | 0 | 0.1315 | 0.4 | 0.66 | 0.074 |
| 0 | 0 | 1 | 0 | 0.1398 | 0.1163 | 0.5644 | 0.4286 |
| 0 | 1 | 0 | 0 | 0.25 | 1.0 | 0 | 0.667 |
| 0 | 1 | 1 | 0 | 0.722 | 0.600 | 0.633 | 0.667 |
| 1 | 0 | 0 | 0 | 0.111 | 0 | 0 | 0 |
| 1 | 0 | 1 | 0 | 0 | 0 | 0.4 | 0.667 |
| 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 |
| | | | 5 | 3 | 1 | 5 | 1 |

obtained there the highest success rate predicting the evasion of the subjects. The hypothesis H_3 require 8 states for representing it. We are interested in predicting the decision of evading at time t therefore in Table 4 we represent the states in eight rows. A number one (zero) in the first column, t-2, indicates that the subject was (not) audited two periods ago. A number one (zero) in the second column, t-1, says that the subject was (not) audited one period ago. A number one (zero) in the third column says that the last time that the subject was audited he was (not) caught evading.

The last five columns represent the empirical probability of evading in every state, for the four different groups. The last row has the representativeness of each of the five automata. For instance automaton Γ_5 represents only one subject.

The last two states, where the subject is audited in period t-1 and t-2 are never visited in our experiments, and this is why all there values are zeros.

There are five subjects, represented by automaton Γ_1 and they never evade. Automaton Γ_4 describes subjects which likely they evade if they were caught the last time that they were audited. This is even more probable if they were just caught in the previous period. Finally automaton Γ_2 represent subjects with low probability of evading. However, if they are caught evading, they with a probability of 0.72 will evade in the next period. The subjects represented by Γ_4 and Γ_2 resemble the type of behavior anticipated for the bomb crater effect.

5 Searching for the best automaton

In the previous section, we use automaton for characterizing the behavior of the subjects in the simulations.

In this section we search for the best automaton, according to the success rate at predicting the extensive behavior of all the subjects in Ξ . In this section we are looking for only one automaton to explain the behavior of subjects.

For accomplishing this goal we run the following three steps algorithm:

1. we propose a hypothesis H together with the automaton Γ_H that describes it.
2. we select a set of experiments $E \subseteq \Xi$ which are the target group of our problem. Notice that if we work with the subjects in experiment e ., the set $E = \{e\}$. If we are interested at obtaining only one automaton to explain the behavior of all the subjects in Ξ , then $E = \Xi$.
3. using an optimization algorithm we search for the values of the vector of probability $P \in \Gamma_H$ that maximize the success rate in E .

Every individual $j \in E$ has a history h_j of evasions. Extensively the history might be represented by a string of binary number $h_j = \{1, 0, 1, \dots, 0\}$, where the first 1 means that the subject evaded in the first period and the zero in the second position means that the subject did not evade in the second period.

Given the automata Γ_H with its the set or probabilities of evading P in every state $s \in S$ it is possible to simulate for a subject j a prediction \hat{h}_j of the decision of evading or not.

Given a set of target experiments $E \subseteq \Xi$, a behavioral hypothesis H , and the automaton Γ_H representing it, we obtain the probabilities $P \in \Gamma_H$ to maximize the success rate for all the subjects in E ,

$$P = \operatorname{argmax}_P \left[\sum_{j \in e \subset E} \chi_j \right], \quad (4)$$

where χ_j is the success ratio of the automaton Γ_H predicting the decision of the subject j .

There are important differences between the techniques applied in the previous section and in this one. The first technique helps in the characterization of agents behavior. The goal of that technique is to realize what provide more information: having more states in the automata or having more automata for discriminate behaviors.

The second technique, is aimed to obtain the automaton that predict better the behavior of the subjects within the experiments and between different groups. In this case we work out a criteria for measuring the quality of an automaton and we look for the best automaton using an optimization algorithm.

5.1 The three state automaton revisited

Suppose again the same hypothesis that we used in the previous section, H_0 : *The decision of evading depends on whether the subject was audited or not in the previous period.*

We use genetic algorithms for searching the optimal vector of probabilities P for every experiment $e \in \Xi$. We present the results in Table 5.

Table 5: Optimal Success Ratio and Probability Vector for Γ_{H_0}

| | Characteristic ^{a b} | e_{11} | e_{12} | e_{21} | e_{22} | e_{31} | e_{32} |
|-------|-------------------------------|----------|----------|----------|----------|----------|----------|
| Chile | <i>Success Ratio</i> | 0.6635 | 0.505 | 0.576 | 0.5478 | 0.526 | 0.611 |
| | Initial State S_0 | 0.2 | 0.53 | 0.43 | 0.36 | 0.62 | 0.21 |
| | Audited State "1" | 0.6631 | 0.361 | 0.0859 | 0.41 | 0.933 | 0.7793 |
| | Not Audited State "0" | 0.022 | 0.80 | 0.0156 | 0.996 | 0.127 | 0.0166 |
| Italy | <i>Success Ratio</i> | 0.555 | 0.52 | 0.699 | 0.739 | 0.5232 | 0.624 |
| | Initial State S_0 | 0.96 | 0.21 | 0.63 | 0.11 | 0.231 | 0.14 |
| | Audited State "1" | 0.95 | 0.821 | 0.717 | 0.168 | 0.1768 | 0.115 |
| | Not Audited State "0" | 0.978 | 0.008 | 0.002 | 0.0117 | 0.124 | 0.6797 |

^aWithin every country, the first row has the success ratio. The second, third and fourth row has the probability of evading in each state.

^bThe reference e_{xy} stands for experiment x, group y

The automaton capture different behavioral patterns. There are six cases where the bomb crater effect is captured by the optimal automaton; three cases for Chile $\{e_{11}, e_{31}, e_{32}\}$ and three for Italy $\{e_{12}, e_{21}, e_{22}\}$. There are two cases representing honest behavior, where the probability of evading in both states is low; one for Chile, e_{21} and one for Italy e_{31} . There are another three cases where the subjects are more likely to evade when they were in the state of not audited; two for Chile $\{e_{12}, e_{22}\}$ and one for Italy e_{32} . And finally there is one experiment where subjects evade systematically, and this is in Italy e_{11} .

Thus basically we can observe four different behavioral patterns as consequence of optimizing the probability vector of the automaton represented by the hypothesis H_0 . The most representative pattern is the bomb crater effect.

5.2 Optimizing the Hypothesis

In this section we use the hypothesis H_0 to H_6 that we used before for searching the optimal automaton for every experiment. We present in Table 6 the hypothesis with the best success ratio for each experiment and group.

Two hypothesis, H_0 and H_1 , obtain very good performance at anticipating the subjects decisions in ten out of twelve cases. Both hypothesis are simple in the sense that the automata that represent them require few states. Both automaton use only information from the previous period. This suggest that the changes during the previous period are enough information for predicting subjects decisions in almost all the experiments

5.3 Does it matter what happens during the first periods?

In the experimental design, every experiment has two groups. The difference between the two groups is that subjects participating in the second group were audited at the very beginning of the experiment. The goal of this design, is to obtain an answer to the question of whether individuals evade less when

Table 6: Hypothesis that explain better the behavior of each experiments with one automaton

| Experiment | Hypothesis | Success Ratio |
|-------------------------------|------------|---------------|
| e_{c1} Base Line, Group 1 | H_0 | 0.6711 |
| e_{c2} Base Line, Group 2 | H_0 | 0.5674 |
| e_{c3} Redistribution, g. 1 | H_1 | 0.5453 |
| e_{c4} Redistribution, g. 2 | H_1 | 0.5904 |
| e_{c5} Public Good, Group 1 | H_3 | 0.6614 |
| e_{c6} Public Good, Group 2 | H_0 | 0.6059 |
| e_{i1} Base Line, Group 1 | H_1 | 0.5967 |
| e_{i2} Base Line, Group 2 | H_2 | 0.6956 |
| e_{i3} Redistribution, g. 1 | H_0 | 0.5287 |
| e_{i4} Redistribution, g. 2 | H_1 | 0.6008 |
| e_{i5} Public Good, Group 1 | H_0 | 0.7384 |
| e_{i6} Public Good, Group 2 | H_0 | 0.6708 |

they are audited at the very beginning of their live.

In order to check for this, we formulate three new hypothesis. This three new hypothesis distinguish whether the subjects were audited or not at the beginning of the experiment.

H_7 Subjects decision of evading depends on whether he was audited or not during the previous period and if he was audited during the first five periods of the experiment he will reduce his probability of evading.

H_8 Subjects decision of evading depends on whether he was audited or not during the previous period and if he was audited during the first five periods of the experiment he will reduce his probability of evading. Moreover the decision of whether to evade or not depends on whether the agent was caught or not the last time that he was audited.

H_9 Subjects decision of evading at time t depends on whether he was audited or not during the two previous period and if he was audited during the first five periods of the experiment he will reduce his probability of evading.

With the new set of hypothesis, from H_0 to H_9 we search for the optimal automaton for anticipating the decision of the subjects for the three experiments in every country. Notice that this time, we do not distinguish between the groups in every experiment. The results are presented in Table 7.

In four experiments the new hypothesis outperform the previous set of hypothesis. This means that there is an effect on auditing the subjects at the beginning of their life as taxpayers.

6 Conclusions

In this paper we use the Theory of Finite Automata for explaining the behavior of subjects in experiments on tax evasion run in Chile and Italy. In every country there were 6 experiments with different

Table 7: Hypothesis that explain better the behavior of each experiments with one automaton

| Experiment | Hypothesis | Success Ratio |
|----------------------------|------------|---------------|
| $e_{c1,c2}$ Base Line | H_7 | 0.5820 |
| $e_{c3,c4}$ Redistribution | H_8 | 0.5924 |
| $e_{c5,c6}$ Public Good | H_7 | 0.5615 |
| $e_{i1,i2}$ Base Line | H_8 | 0.5810 |
| $e_{i3,i4}$ Redistribution | H_0 | 0.7186 |
| $e_{i5,i6}$ Public Good | H_0 | 0.6048 |

designs. We are interested at the decisions made by the individuals about whether to evade or not. Every automaton represent an hypothesis about how subjects make their decision about evasion.

We approached the experimental data using a Moore Automata with two different goals. On the one hand we try to identify classes of subjects behaviors within an experiment and on the other hand we search for the automaton, within every experiment, with the highest success rate at predicting subjects decisions.

Related to the first goal, we identify two patterns that explain more than 70% of subjects behaviors: there are subjects that never evade and subject that evade strategically. The subjects evading strategically behave following the bomb crater effect mentioned in [1]. They mainly evade after they have received a tax control. This strategic agent make up their mind to think that the probability of receiving two tax control in a row is very low.

We also discover that a simple behavioral hypothesis explains a big proportion of subject decisions. Indeed, if we move into more complex hypothesis the success rate is lower than simple ones. The complexity in this case is defined by the number of states that an automaton require for capturing the hypothesis: more complexity means more states.

Related to the second goal, when we search for an automaton explaining the highest proportion of subject decisions, the success ratio depends on the experiments and on the type of decisions made by the agent within the experiments. When subjects in an experiment rarely decide to evade, it is easy to anticipate their decisions. Thus, the success rate is high. When subjects decisions are less biased, the success ratio decrease.

We showed that the Theory of Finite Automata might be an interesting tool for understanding subjects behavior in the experiments. In future work we hope to integrate the knowledge developed in this paper for simulating the behavior of economic agents.

References

- [1] L. Mittone, “Dynamic behaviour in tax evasion: An experimental approach,” *The Journal of Socio-Economics*, vol. 35, pp. 813–835, 2006.

- [2] A. Rubinstein, “Finite automata play the repeated prisoner’s dilemma,” *Journal of Economic Theory*, vol. 39, pp. 83–96, 1986.
- [3] M. Allingham and A. Sandmo, “Income tax evasion: a theoretical analysis,” *Journal of Public Economics*, vol. 1, pp. 323–338, 1972.
- [4] S. Yitzhaki, “A note on income tax evasion: A theoretical analysis,” *Journal of Public Economics*, no. 3, pp. 201–202, 1974.
- [5] E. Kirchler, *The Economic Psychology of Tax Behaviour*. Cambridge University Press, 2007.
- [6] D. Kahneman and A. Tversky, “Prospect theory: An analysis of decision under risk,” *Econometrica*, vol. XLVII, pp. 263–291, 1979.
- [7] M. E. Romera, “Using finite automata to represent mental models,” master of arts, San Jos State University, August 2000.
- [8] R. Selten, “Aspiration adaptation theory,” *Journal of Mathematical Psychology*, vol. 42, pp. 191–214, 1998.