

Effects of inspections in small world social networks with different contagion rules

Francisco Muñoz^a, Juan Carlos Nuño^{a,*} and Mario Primicerio^b

December 10, 2014

^a Department of Applied Mathematics. Universidad Politécnica de Madrid, 28040-Madrid, Spain

^b Dipartimento di Matematica “Ulisse Dini”. Università degli Studi di Firenze, 55015 Firenze, Italy

(*) Corresponding author:

e-mail address: juancarlos.nuno@upm.es (Juan Carlos Nuño)

Abstract

We study the way the structure of social links determines the effects of random inspections on a population formed by two types of individuals, e.g. tax-payers and tax-evaders (free riders). It is assumed that inspections occur in a larger scale than the population relaxation time and, therefore, a unique initial inspection is performed on a population that is completely formed by tax-evaders. Besides, the inspected tax-evaders become tax-payers forever. The social network is modeled as a Watts-Strogatz Small World whose topology can be tuned in terms of a parameter $p \in [0, 1]$ from regular ($p = 0$) to random ($p = 1$). Two local contagion rules are considered: (i) a continuous one that takes the proportion of neighbors to determine the next status of an individual (node) and (ii) a discontinuous (threshold rule) that assumes a minimum number of neighbors to modify the current state. In the former case, irrespective of the inspection intensity ν , the equilibrium population is always formed by tax-payers. In the mean field approach, we obtain the characteristic time of

convergence as a function of ν and p . For the threshold contagion rule, we show that the response of the population to the intensity of inspections ν is a function of the structure of the social network p and the willingness of the individuals to change their state, r . It is shown that sharp transitions occur at critical values of ν that depends on p and r . We discuss these results within the context of tax evasion and fraud where the strategies of inspection could be of major relevance.

Keywords: tax-evasion, social network; mathematical model; computer simulations

Classification codes: C6, C7

1 Introduction

In the past few years mathematical and computer modelling have been used to provide new and relevant insights in the mechanisms of diffusion of fraud and tax evasion in modern societies [19, 32, 31, 4, 1] thus, suggesting how to optimize strategies to contrast such phenomena. A particular attention has been focused on how unlawful behaviour can induce a sort of contagion on the "neighbouring" individuals in a given society. Indeed, the personal attitude facing social compliance depends very much on the opinion of other partners (citizens) and, in a global scale, on public opinion [11]. The way our neighbors, e.g. friends, colleagues or, in general, peers, face the social contract highly affects our position to this respect. Indeed, the stability of human groups depends on the principle of cooperation that is implemented, among other factors, in tax compliance [13]. When cooperation is not integral, i.e. when some participants do not contribute to the common pool then, group coexistence becomes fragile.

Interventions are widely applied to control network behavior [27]. In particular, inspections and, if necessary, punishment is an effective way of fighting against tax evasion. Inspections are costly and, therefore, they have to be implemented in a rational way [22]. A complete inspection, i.e. to every citizen of the society, becomes unpractical even for small size populations. Instead, random inspections are feasible in general. The question arises about which is the optimal intensity, i.e. number of inspected individuals, to be applied.

Moreover, the timing of inspections could become relevant in certain situations. These decisions are obviously taken under a limited budget scenario that complicates even more the optimal solution of the problem. In this paper we assume that a unique set of inspections is initially performed to the whole population. The intensity of the inspections, i.e. the size of the set of individuals that is inspected, is taken as a control parameter of the problem. Consequently, we study the equilibrium distribution of the population. This is consistent with assuming that inspections occur at a larger time scale than the relaxation period of the population. For simplicity, we consider that individuals can only take two states: law-abiding (tax payers) and free-riders (tax-evaders). Human societies are linked, i.e. individuals are not isolated but they relate to each other by social links that transmit, for instance, their propensity to pay taxes. The way the social contagion occurs and the structure of the social network are two main aspects that determine the final outcome of inspections. An individual can belong to different social networks, e.g. professional or private, and the structure of these networks can be very different [28]. For instance, the social network of scientists is proven to have a scale free topology and the same topology appears in the network formed by sexual contacts among individuals (see, for instance, [5] for a review). However, friendship and peer networks seem to exhibit a small world structure [30]. Information concerning tax evasion is a delicate matter and consequently, it tends to flow through networks of small world type. Therefore, we will focus on them in this work.

A widely applied assumption treats social contagion from an ecological perspective as an epidemic, where the probability of an individual (i) to be infected at time t , $H(i; t)$, is proportional to the number of its infected neighbors at that time, $N(i; t)_I$:

$$H(i; t) = \lambda N(i; t)_I$$

where $\lambda > 0$ is a contagion rate. However, recent investigations have shown that the micro-processes involved in social contagion are complex [6, 25, 7, 3]. A family of models take into account social peculiarities by assuming a discontinuous response of a node to the inflow information [24]. These models are inspired in the classical threshold model[12] where the adoption of an initiative depends basically on whether a minimum number of neighbors have already adopted it.

Recent works have shown that social diffusion in real networks are of this type [8, 2, 18]. Nonetheless, in other cases, probabilistic diffusion mechanisms seem to describe better the behavior of the network [10]. Besides, the individuals personality is known to play an important role in the adoption of certain features [23]. For this reason, in our model we assign a personality to the nodes, measured by a parameter r , that drives the willingness to change their state.

The remainder of the paper is organized as follows: Next section presents in detail a simple model and defines its main parameters. Sections 3 and 4 study the effects of inspections on the population for different network topology for two different contagion rules. We discuss the implications of our results in the last section.

2 A toy model

Let us assume a population of interacting individuals that form a connected social network. We assume that the network is homogeneous that is, all the nodes have approximately the same number of neighbors (equal average connectivity). In other words, the network has a characteristic scale with regards to its connectivity. In order to tune the degree of randomness we take the Small-World model of Watts and Strogatz (WS) [29] as the reference network. This family of networks has been identified with real social networks because its short path length (the diameter of the network increases logarithmically with the number of nodes) and the high clustering coefficient. A regular ring of N nodes symmetrically connected with k nearest neighbors is transformed by rewiring the links to randomly chosen nodes with a probability p . By construction, the WS network lies in between two limit types of networks: regular (ring) and random [29]. If $p = 0$ the regular network is preserved, whereas for $p = 1$ a random network is formed. WS networks have connectivity distributions $P(k)$ that are peaked at an average value $\hat{k} = \langle k_i \rangle$ and decaying exponentially for $k > \hat{k}$ and $k < \hat{k}$ (poissonian like distribution).

The influence of inspections in the time evolution of the population depends on (i) the intensity applied (ν), (ii) the network topology (p) and (iii) the local diffusion (updating) rules (e.g. continuous or discontinuous). The inspections reduce the number of nodes that can become evaders at any time. In all cases, for

each intensity of the inspections, the population evolves towards an equilibrium state characterized by a distribution of evaders and payers. The time it takes the population to converge depends on the parameters. For the parameter setup used in the simulations, it is observed that after 10000 time steps the stationarity of the population is assured. The final estimation of the equilibrium population is obtained from the average of 25 simulations of each case. In all the cases, the total population is fixed to $N = 5000$. This yields a network of N nodes. The networks have been created using the Python Graph Library Networkx (*networkx.github.io*) [16] following the Watts-Strogatz Small World model [29]. The number of links K is also kept during the time evolution. Most of the results have been obtained for $K = 25000$, that corresponds to an average connectivity $\hat{k} = 5$. Nonetheless, other values have been applied in specific cases. In all cases, the initial population is formed entirely by evaders. Note that this corresponds with the less favorable situation to fight tax-evasion.

Due to the social interactions, information can flow through the network nodes. It is assumed that this information restricts the state a node has at each time. As it has been already stated, a node (individual) can adopt only two states: tax payer or tax evader. At each time step, a node is influenced by the current state of its nearest neighbors. Different updating (contagion) rules can affect considerably the state of the nodes at the next time. In particular, two specific mechanisms are considered: Continuous Rule and Discontinuous Threshold Rule.

3 Continuous rule

According to this rule, the probability of changing the state of node i with connectivity $K(i)$ from state S_2 to state S_1 is proportional to the ratio $N(i)_{S_1}/K(i)$, where $N(i)_{S_1}$ is the number of neighbors that are in state S_1 at this time. Formally,

$$P(i)_{S_2 \rightarrow S_1} = \frac{N(i)_{S_1}}{K(i)} \quad (1)$$

Obviously, the probability of remaining in the current state S_2 is:

$$P(i)_{S_2 \rightarrow S_2} = 1 - P(i)_{S_2 \rightarrow S_1}$$

A similar definition is applied for the probability of change from S_1 to S_2 .

The application of inspections, irrespective of the intensity, causes that the whole population becomes tax-payers. In other words, the population of evaders tends to 0 with probability 1 irrespective of the number of nodes initially inspected. The time the population takes to achieve this equilibrium population formed completely by payers depends on the parameter p , the intensity of the inspections and the average connectivity. Since the average path length of the WS networks decreases with p , it is expected that this characteristic time also decreases with p . This can be confirmed by studying the mean-field equation of the system.

It is well known from the theory of transmissions of infectious disease that continuous local contagion rules give rise to average diffusion terms of "bimolecular type", in the sense that the rate of transition from the state of "susceptible" to the state of "infectious" is proportional to the product of the number of members of both classes (see, for instance, [21]). Besides, the Small World networks of Watts and Strogatz are homogeneous, i.e. they have a unique scale of connectivity. Since the statistical behaviour of all the nodes is similar, at a mean-field level, the population dynamics can be described by a single equation.

For any given t , let $x(t)$ and $y(t)$ be the density of payers and evaders respectively. It is assumed that initially the whole population is formed by tax evaders but that a given fraction ν of nodes is inspected and that such nodes remain tax payers forever, so that only a fraction $1 - \nu$ of nodes change their state as time passes and at each time the normalization condition $x(t) + y(t) = 1 - \nu$ is satisfied.

Let us start by considering the simplest case of an homogeneous network in which each node is connected to all the remaining nodes ($k = N$). At each time the evolution of the fraction of tax-payers $z(t) = x(t) + \nu$ is described by a balance equation that expresses the fact that the rate of change of the number of tax payers is the difference between the rate at which tax evaders that are converted into tax payers and the rate at which tax payers are converted in evaders. According to the continuous "contagion" rule we have the following differential equation:

$$\frac{dz}{dt}(t) = \alpha [z(t)(1 - z(t)) - (1 - z(t))(z(t) - \nu)] \quad (2)$$

where α is a constant representing the "time scale" for the evolution. Thus, we

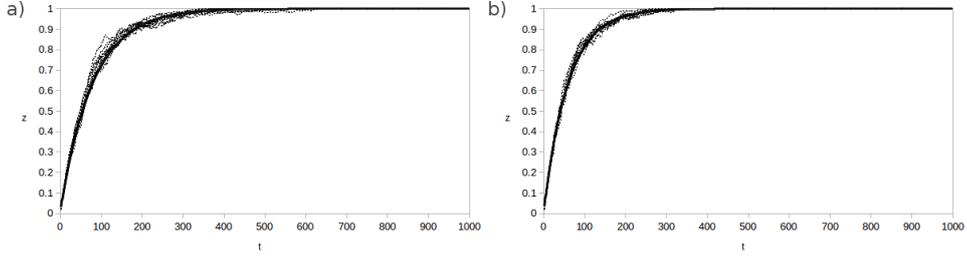


Figure 1: Time evolution of the proportion of payers z obtained from 10 simulations and the corresponding continuous solution (5) of the mean field differential equation (4) with a value of Θ_0 given by equation (7) for $\hat{k} = 5$ and $\nu = 0.02$. The values that define the Small-Network topology are: $p = 0.1$ and $p = 1$. The characteristic time of the population is given by equation (6) and depends on p . For instance, it can be easily calculated that $T_c \approx 77.52$ for $p = 0.1$.

have the mean field equation:

$$\frac{dz}{dt}(t) = \alpha \nu (1 - z(t)) \quad (3)$$

In case of general networks, it is reasonable to start assuming that the "time scale" depends on both the average connectivity \hat{k} and the parameter p and hence, to write:

$$\frac{dz}{dt}(t) = \Theta_0(\hat{k}, p) \nu (1 - z(t)) \quad (4)$$

For instance, in the trivial case $k = 0$ in which each node is isolated we would have $\Theta_0(0, p) = 0$. Starting from the initial condition $z(0) = \nu$ we get:

$$z(t) = 1 - (1 - \nu) e^{-\nu \Theta_0(\hat{k}, p) t} \quad (5)$$

For $\hat{k} > 0$, the unique equilibrium point is $z = 1$ irrespectively of the value of ν , meaning that the asymptotic population is formed exclusively by payers. The main difference with respect to the case of propagation of epidemics (in which, in the simplest case, the solution is a logistic curve) consists in the asymmetry induced by the fact that, in case of infectious diseases the contagion acts just from the infectious individuals to the susceptible ones and not vice-versa. The rate at which the equilibrium point is approximated depends on the fraction of the inspected nodes and on the structure of the network. Indeed, although the time to achieve the equilibrium is infinite, it is still possible to get an estimate of how quickly the asymptotic converge occurs. This estimate is usually provided

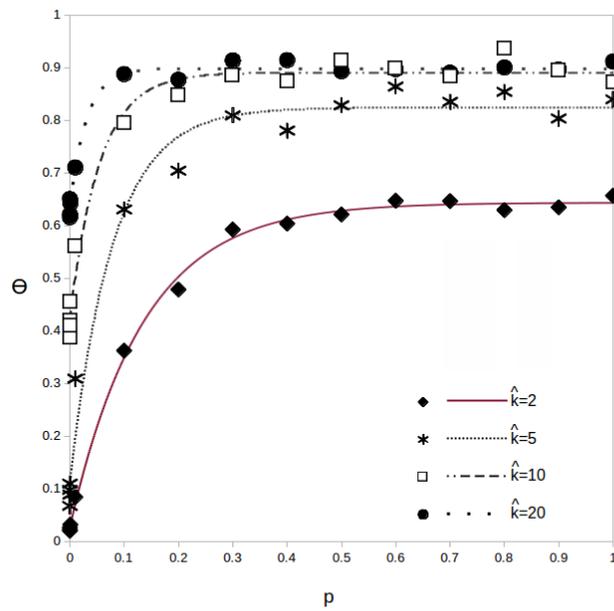


Figure 2: Fit of the diffusion rate Θ_0 in terms of p for some values of \hat{k} : 2 (diamonds and continuous line), 5 (stars and dashed line), 10 (squares and dash-point line) and 20 (circles and point line). The total population, i.e. the number of nodes is $N = 5000$. For each simulation we get a value of Θ_0 . Each point of the graph is obtained from the average of ten Θ_0 -values. As it can be seen, the agreement between the numerical data and the fitting curves is good.

Table 1: Estimated value of the parameters of function (7) for different values of \hat{k} .

\hat{k}	α	β	γ	R-square
2	0.6435	0.6163	7.388	0.9982
5	0.8241	0.7126	12.79	0.9839
10	0.8906	0.4643	18.16	0.9845
20	0.8987	0.2679	32.71	0.9906

by the value of the "characteristic time" [20], denoted as T_c , which in the cases where Θ_0 is constant, e.g. for large values of p and \hat{k} , is given by the inverse of the absolute value of the eigenvalue of equation (4) :

$$T_c = \frac{1}{\nu \Theta_0(\hat{k}, p)} \quad (6)$$

Concerning the dependence of Θ_0 on \hat{k} and p , it can be inferred that it is monotonically dependent on both variables. Moreover, if \hat{k} is fixed, it is natural to assume that the derivative $\frac{d\Theta_0}{dp}$ tends rapidly to zero as p increases. Conversely, if p is fixed, Θ_0 goes from 0 (for $\hat{k} = 0$) to an asymptotic value which is slightly increasing for increasing p . The functional dependence of Θ_0 on \hat{k} and p can only be determined from the simulations. The comparison of the numerical data with the solutions of (4) for different values of \hat{k} and p allows to obtain a good fitting with the function:

$$\Theta_0(\hat{k}, p) = \alpha(\hat{k}) - \beta(\hat{k}) e^{-\gamma(\hat{k})p} \quad (7)$$

where some values of $\alpha(\hat{k})$, $\beta(\hat{k})$, $\gamma(\hat{k})$ are given in table 1. Importantly enough, the R -square of the fit is above 0.98 in all the considered cases. As it can be seen in Fig.1, Θ_0 increases very rapidly with p for low values of p and it saturates for $p > 0.2$. The saturation value depends on the average connectivity \hat{k} and increases with it. This value is almost constant for any $\hat{k} > 10$ thus proving that, in practice, the approximation (2) that we postulated for $\hat{k} = N$ is valid for any "sufficiently connected" network. In Fig. 1 we show the time evolution of payers $z(t)$ obtained from simulations and by the mean field approximation for large enough values of p . The agreement is excellent for all times and for all values of p between $p = 0.1$ and $p = 1$.

Remark 3.1

It has to be noted that the agreement between the mean field equation (4) and the simulations is not adequate for lower p -values when the network becomes almost regular, more precisely for $p < 0.1$. Note that for the same parameter setup used in Fig. 1, i.e. a network of size $N = 5000$, $\hat{k} = 5$ and $\nu = 0.02$, the number of initial payers (infectious agents) is very low, approximately 100, that will be connected with low probability. This implies that initially, the contagion rate of this behavior according with the probabilistic rule (1) will be extraordinarily low. Therefore, it is expected that in these initial times, an "anomalous" diffusion occurs [26, 9].

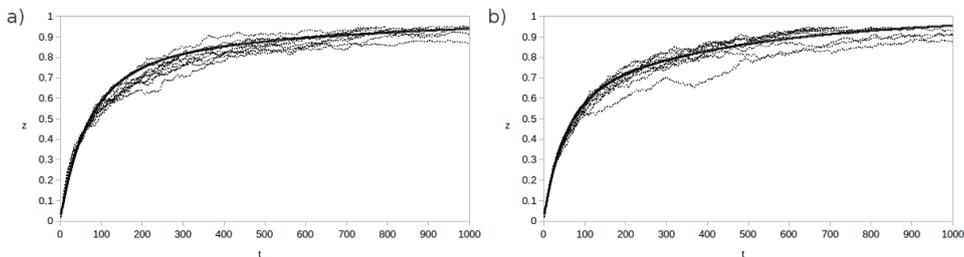


Figure 3: Time evolution of the proportion of payers z obtained from 10 simulations for $p = 0$ and $p = 0.0001$. For these low values of p we obtain a adequate fit using the mean field approximation described by equation (8) with a time-dependent diffusion rate. The function Θ is assumed to change accordingly to equation (9) with the fitting parameters given in Table 2. As in the previous figure, $\hat{k} = 5$ and $\nu = 0.02$.

This means that a more accurate description of the "contagion" by a mean field approximation can be obtained if we substitute equation (4) by the non-autonomous equation [14, 17]:

$$\frac{dz}{dt}(t) = \Theta(\hat{k}, p; t) \nu (1 - z(t)) \quad (8)$$

where the function Θ differs from the function Θ_0 just during the transient period. Figure 3 depicts the time evolution of the population of payers obtained from the simulations and from non-autonomous mean field equation (8). As it can be seen the agreement is good.

We have fitted Θ with simulations using the simple form:

$$\Theta(\hat{k}, p; t) = a(\hat{k}, p) e^{-b(\hat{k}, p)t} + \Theta_0(\hat{k}, p) \quad (9)$$

where the values of parameters a, b and Θ_0 can be found by comparison with simulations. Note that the first term is significant only at initial times. This transitory period depends on p and \hat{k} and tends to 0 when p tends to 1 and \hat{k} increases. To our knowledge, there is not a complete theory about the spread of information in complex networks in extreme situations as those occurred at initial times in our model when the initial "infectious nodes" are connected with very low probability. Even the classic mean field approach is questioned in these situations. We think that this subject is out of the scope of this work and it will be discussed with some more detail in a forthcoming paper.

4 Discontinuous threshold rule

As stated in the introduction, local diffusion in social networks turns out to be more complex than the continuous rule studied in the previous section. In particular, threshold rules seem to mediate actively in the contagion of a large group of social processes. To take into account these cases, we define a willingness of node i to keep its current state S_1 (payer or evader) as:

$$W(i)_{S_1 \rightarrow S_1} = \frac{N(i)_{S_1}}{K(i)} \quad (10)$$

Similarly, the willingness to adopt the state S_2 is given by:

$$W(i)_{S_1 \rightarrow S_2} = \frac{N(i)_{S_2}}{K(i)} \quad (11)$$

As before, $N(i)_{S_1}$ and $N(i)_{S_2}$ are the number of neighbors of node i that are in the states S_1 and S_2 , respectively and $K(i) > 0$ is the connectivity of node i . Note that $W(i)_{S_1 \rightarrow S_2} = 1 - W(i)_{S_1 \rightarrow S_1}$. The final decision of node i under the influence of its neighborhood depends on its personality that we take into account by defining a parameter $r \in [0, \frac{1}{2}]$, such that:

- (i) The node i at state S_1 will remain in this state if $W(i)_{S_1 \rightarrow S_1} > r$ and $W(i)_{S_1 \rightarrow S_2} < r$.
- (ii) Node i will change to S_2 if $W(i)_{S_1 \rightarrow S_1} < r$ and $W(i)_{S_1 \rightarrow S_2} > r$.
- (iii) The state of node i at the next time is taken randomly between S_1 and S_2 if $W(i)_{S_1 \rightarrow S_1} \geq r$ and $W(i)_{S_1 \rightarrow S_2} \geq r$.

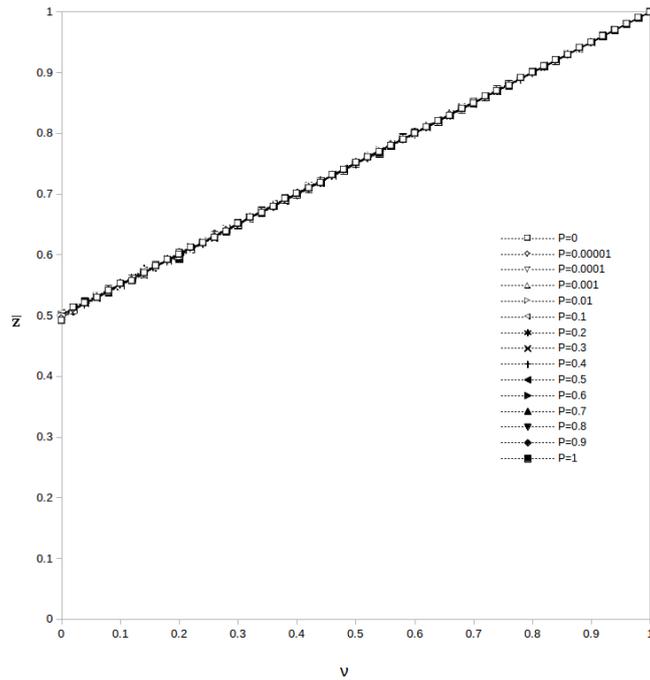


Figure 4: Equilibrium density of tax-payers \bar{z} as a function of the intensity of the initial inspections ν for different values of the network parameter p . A threshold contagion rule is used with a node personality given by $r = 0.0$. As before, the total population is $N = 5000$. In all simulations the initial population is formed entirely by tax-evaders. After the inspections, a fraction ν of tax-payers is initially present. All the points correspond to the average over 10 realizations taken after 10^4 time steps. Due to the lack of personality the equilibrium population of payers increases linearly with ν , from $z = \frac{1}{2}$ for $\nu = 0$ to $z = 1$ when all the nodes are initially inspected, i.e. $\nu = 1$

Note that, the corresponding cases for $\frac{1}{2} < r \leq 1$ can be obtained by symmetry from the ν -values ranging in the first half unit interval. As in the previous model, those nodes that have been inspected adopt the state of payers forever.

The consideration of this threshold contagion rule modifies drastically the asymptotic outcome of the population. The simplest situation is given by the limit case $r = 0$ that corresponds to a population totally formed by indecisive individuals. For this r -value, irrespective of the current state, a node chooses randomly its next state. As Fig. 4 shows, the equilibrium population of payers \bar{z} increases linearly with ν with a slope $\frac{1}{2}$. This means that, independently of p (and also of \hat{k}), half of the remaining population is converted to payers. Nonetheless, the values of these parameters p and \hat{k} , affect the way the equilibrium population is attained.

The situation is more complicated for $r > 0$. For small intensity of inspections, the equilibrium population contains tax-evaders. The value of this proportion depends on the intensity ν , the structure of the network p and also, on the value of r . Figure 5 represents the equilibrium proportion of payers as a function of the intensity of the inspections ν for different values of p when $r = 0.2$ and the average connectivity is $\hat{k} = 5$. As it can be seen, three regimes appear: for low values of ν , the population is almost fully formed by evaders, irrespective of the value of p . This forms an upper branch in the bifurcation diagram. If $p > 0.1$ then, the population remains stable in two intermediate levels; a first one until the intensity of the inspections is lower than $\nu \approx 0.04$. A second intermediate level, that corresponds to approximately half of the whole population, exists for $0.04 < \nu < h(p)$, where h is a function of p . For $\nu > h(p)$ the equilibrium population becomes tax-payer. As it can be observed, if $\nu > 0.2$ then, the asymptotic population is formed by tax-payers for all values of p .

The intermediate branch disappears when $r = 0.5$, i.e. if the threshold rule is a pure majority rule (see Fig. 6). In this case, the interval of uncertainty where $W_{S_1} > r$ and $W_{S_2} > r$ does not exist and, consequently this coexistence branch is no longer present in the bifurcation diagram. The equilibrium population of evaders decreases monotonically as the intensity of the inspections increases until ν reaches a critical value that abruptly precipitates the evader asymptotic population. These critical values depend on p as depicted in Fig. 6. Its exact

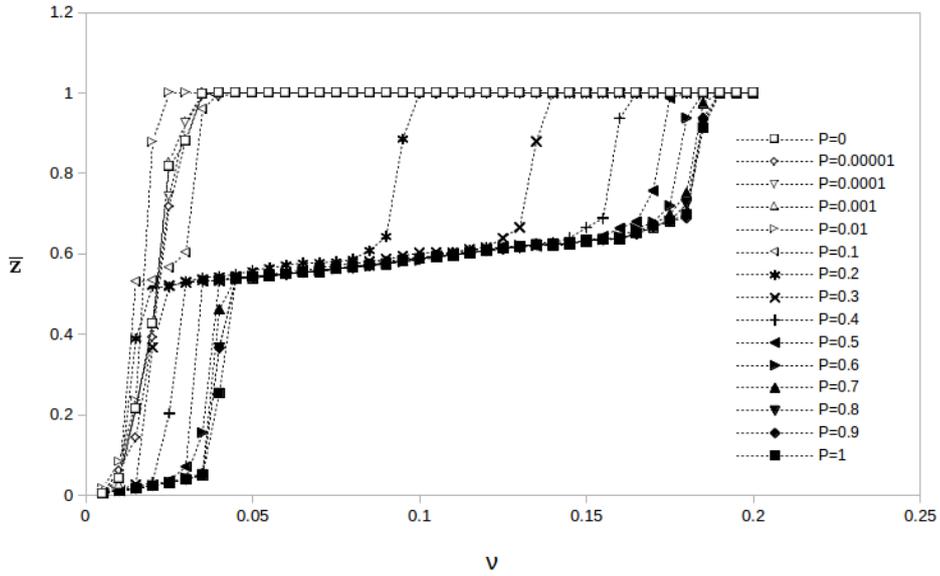


Figure 5: As in the previous figure, the equilibrium value \bar{z} is plotted as a function of ν for different values of p . The value of the node personality $r = 0.2$. As before, the initial population is formed entirely by tax-evaders and, after inspections, a fraction ν of tax-payers is generated that remains forever. All the points correspond to the average over 25 realizations taken after 10^4 time steps. As it can be observed, an intermediate equilibrium population appears for $0.1 \leq p \leq 1$ that grows linearly with ν . At this level, the population is equally divided between payers and evaders. It is a consequence of the contagion threshold rule that assumes a random choice when there is not a priority to either stay or change the current state of an individual (node) for the case $r = 0.2$. Compare this situation with that depicted in Fig.6 for $r = 0.5$, i.e. a majority rule.

value can be approximately estimated from the maximum of the corresponding standard deviation of the population for each p -value.

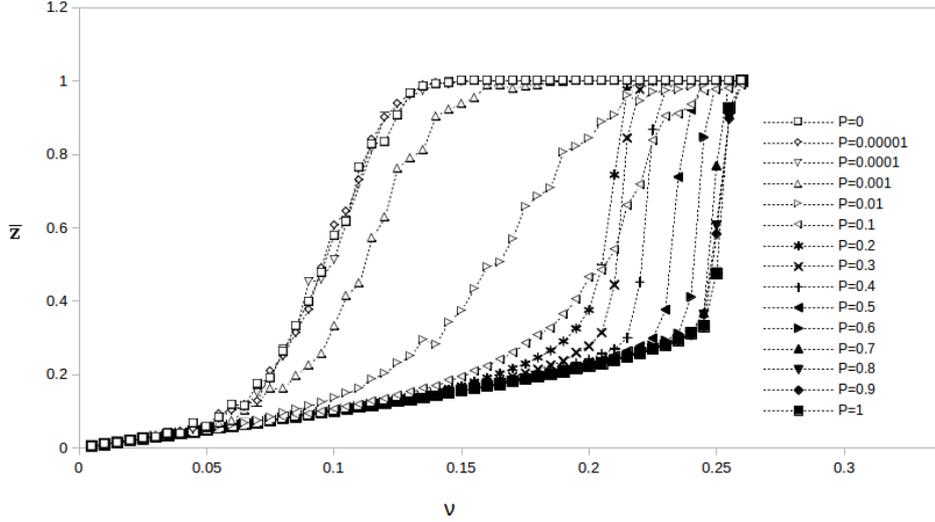


Figure 6: As in Fig. 5 but with a personality $r = 0.5$. Now, contrary to what occurs in the previous figure 5 no intermediate branch appear for any value of ν . As in the case $r = 0.2$ lower values of p are more sensitive to inspections, i.e. the equilibrium population of payers achieve the upper branch for lower values of ν

For the limit $\hat{k} = N$, i.e. all nodes interact with each other, an analytical solution of the mean field equation can be found (see Appendix 1). From this solution the asymptotic behavior of the population in correspondence of any pair (ν, r) can be derived (see Fig. 6). Whether a mean-field approximation for any \hat{k} could be written in terms of some approximation of the Heaviside functions, depending on \hat{k} and p as in the continuous contagion rule is under investigation.

A "mixed" rule, between the continuous rule and the discontinuous threshold rule can also be considered. Two possibilities come to mind:

- (1) Changing the rule expressed under (iii) of Sec.4 and assuming instead that if $W(i)_{S_1 \rightarrow S_1} \geq r$ and $W(i)_{S_1 \rightarrow S_2} \geq r$, i.e. the uncertainty interval, then the state of node i will evolve according to probabilities $P(i)$ that coincide with $W(i)$.
- (2) Let us denote by $C(S_1 \rightarrow S_2)$ the probability of passing from state 1 to state

2 according to the continuous rule and by $T(S_1 \rightarrow S_2)$ the probability of passing from state 1 to state 2 according to the threshold rule. Then, it could be assumed that

$$P(S_1 \rightarrow S_2) = cC(S_1 \rightarrow S_2) + (1 - c)T(S_1 \rightarrow S_2) \quad (12)$$

for a fixed value of $c \in (0, 1)$.

In the Appendix 2 we present the solution of the mean field approach that corresponds to the mixed rule (1) for the limit case $k = N$. It can be appreciated that, when compared with the discontinuous rule, the behaviour in this case is simpler due to the absence of the uncertainty region (see Figures 7 and 8).

5 Discussion

The existence of interactions among the members of a society brings about a social network where the nodes are individuals and the edges (links) represent the relationships among nodes. The way a social network reacts to external influences, e.g. inspections, depend mainly on their structure and the way information is transmitted through the network, i.e. social contagion. In this paper we have studied a toy model that considers that the individuals of a population form a small world social network. We use the model of Watts and Strogatz to tune the degree of randomness of the network from regular ($p = 0$) to pure random network ($p = 1$). For each value of p , we have applied an inspection strategy that consists in converting a fraction ν of nodes that are tax-evaders into tax-payers. We assume that inspected nodes are tax-payers forever in the time scale of the population dynamics. We have also considered two local contagion rules: continuous and discontinuous. For each parameter setup we have applied both rules and we have confirmed that completely different results are obtained. The equilibrium population for the continuous probabilistic contagion rule is formed only by tax-payers, irrespectively of the intensity ν of the initial inspections and the structure of the network p . In contrast, when the threshold rule is applied, the equilibrium population depends on both ν and p . Besides, different willingness to change under the threshold rule gives rise to distinct bifurcation diagrams as depicted in Figures 5 and 6. For the continuous contagion rule, we have applied a mean field approach to relate the

growth rate Θ with p and the average connectivity \hat{k} . It has been proven that the characteristic time to achieve the equilibrium population decreases inversely with the intensity of inspection and increases with p and \hat{k} that, ultimately, are a measure of the characteristic path length of the network.

These results offer new alternatives to the authorities that fight tax evasion and fraud. Knowing the topological properties of the social network allows to estimate the network parameters p and \hat{k} , assuming to be of small world type. With this knowledge, it is possible to apply an intensity of inspections to achieve the optimal response with the lowest cost. For instance, if we discover that the structure of the social network is compatible with a small world of Watts-Strogatz with $p = 0.2$ and $\hat{k} = 5$ and we assume a threshold contagion rule with a willingness to change of $r = 0.2$ then, the minimum intensity of inspections that assure the lowest population of free riders is $\nu = 0.1$. If the willingness to change is $r = 0.5$, instead the intensity of inspections that assure a lowest cost is $\nu = 0.23$. But even if in case of not having a complete certitude about the structure of the network, the model provides a largest value of the intensity of the inspections to get a complete eradication of free-riders, around $\nu = 0.19$ for $r = 0.2$ and $\nu = 0.26$ for $r = 0.5$. In the extreme case $r = 0$, a non-null population of evaders remains for every $\nu < 1$ (see Fig. 5). The absence of personality requires greater efforts to fight tax-evasion.

We would like to end this discussion by making a short reference to the effect of the choice of the inspected individuals in the population. In our model, all the inspections have been carried out randomly, i.e. the nodes to be inspected have been selected without any additional information. However, when the resources are limited, the choice of the correct nodes to be inspected is a critical issue of major relevance [27, 15]. It is true that when the network is homogeneous, as the Small-World model, the preference in the inspection attending to the node degree is not relevant since all the networks have the same average connectivity. On the contrary, for inhomogeneous networks the way inspections are performed could be decisive. Indeed, a preferential vaccination based on the connectivity degree has been proven very useful to control the spread of epidemics on scale free networks [21].

Acknowledgements

Francisco Muñoz thanks the Mediterranean Office for Youth Program by the mobility grant (MOY GRANT N+2010/024/01) in the Program Master in Mathematics and Mathematical engineering. Juan Carlos Nuño wishes to acknowledge the European Erasmus staff mobility programmes. We are grateful to Laura Navajas for checking the English of this manuscript. We also thank the reviewers for their useful suggestions and remarks.

Appendix 1: Mean field approximation for the threshold rule

For the limit case $\hat{k} = N$, i.e. all nodes interact with each other,

$$W(i)_{X \rightarrow Y} = y; \quad W(i)_{Y \rightarrow X} = x + \nu$$

we normalize the diffusion coefficient by changing the time scale and we consider the different situations that may arise separately.

Let us assume that $0 < r < \frac{1}{2}$ and the initial condition $z(0) = \nu$. For a given value of p the solution as a function of the initial condition ν and the value of r reads:

(I) If $\nu < r$ then, equation (8) reduces to:

$$\frac{dz}{dt}(t) = \nu - z(t) \tag{13}$$

The solution of the corresponding Initial Value Problem (IVP) is : $z(t) = \nu$.

(II) If $r < \nu < 1 - 2r$ then,

$$\frac{dz}{dt}(t) = -z(t) + \frac{1}{2}(1 + \nu) \tag{14}$$

In this case, the solution of the IVP is given by:

$$z(t) = \frac{1}{2}((1 + \nu) + (\nu - 1)e^{-t})$$

that tends asymptotically to the equilibrium solution: $\bar{z} = \frac{1}{2}(1 + \nu)$.

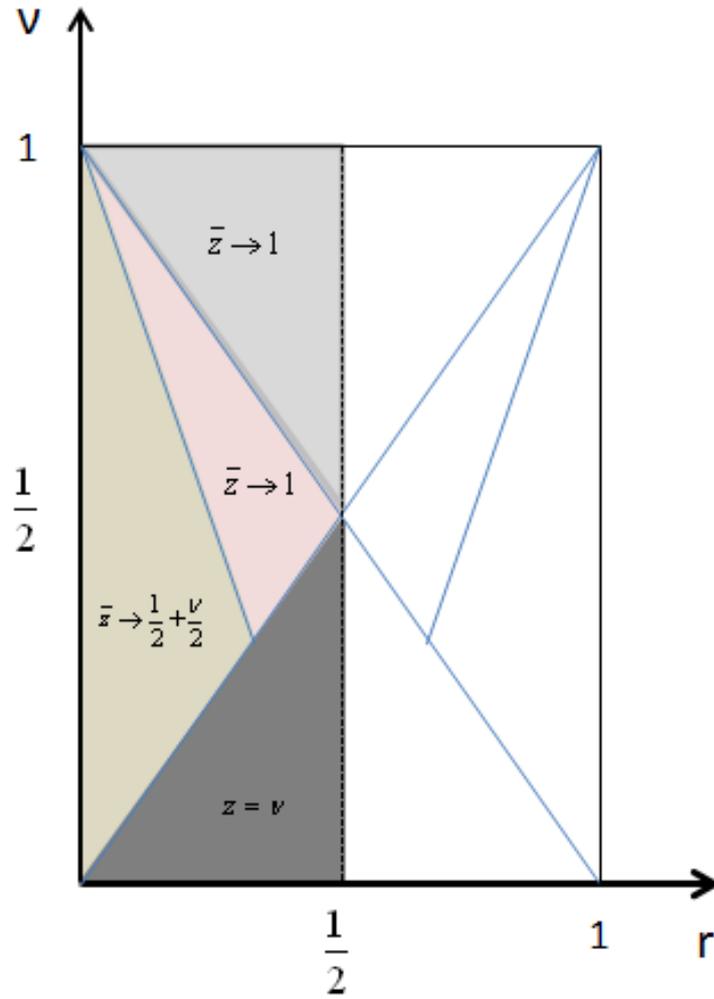


Figure 7: Asymptotic behaviour of the population of payers as a function of the initial condition $z(0) = \nu$ and the parameter $r \leq \frac{1}{2}$ for a WS network with $\hat{k} = N$. Note that the behaviour is symmetric with respect to the line $r = \frac{1}{2}$.

(III) If $\nu > 1 - r$ then,

$$\frac{dz}{dt}(t) = 1 - z \quad (15)$$

and the solution of the IVP is:

$$z(t) = 1 + (\nu - 1)e^{-t}$$

that approaches the equilibrium value $\bar{z} = 1$ as t tends to infinity.

(IV) If $1 - 2r < \nu < 1 - r$ then, the population starts in a situation of uncertainty and, as in (II), the population is described by equation (14). After some time, it switches to case (III) and the population is described by (15). Consequently, the equilibrium value for this case coincides with (III), i.e. $\bar{z} = 1$

Appendix 2: Mean field approximation for the "mixed" rule

In this section, we present the analytical solution of the mean field equation when $\hat{k} = N$ for the mixed rule that considers a probabilistic transition in the uncertainty interval, as described at the end of section 4.

In this case, the differential equations that result as a function of the relation between the initial condition ν and r are:

1. If $\nu > r$ and $1 - \nu < r$ then,

$$\frac{dz}{dt}(t) = 1 - z(t)$$

2. If $\nu < r$ and $1 - \nu > r$ then,

$$\frac{dz}{dt}(t) = -(z(t) - \nu)$$

3. if $\nu > r$ and $1 - \nu > r$ as well as if $\nu < r$ and $1 - \nu < r$ then,

$$\frac{dz}{dt}(t) = z(t)(1 - z(t)) - (1 - z(t))(z(t) - \nu) = \nu(1 - z(t)) \quad (16)$$

This means that, if we start from ν that is in the uncertainty region then, ν is in the interval $(r, 1 - r)$, assuming $r < \frac{1}{2}$. In this case, the ODE for z is given,

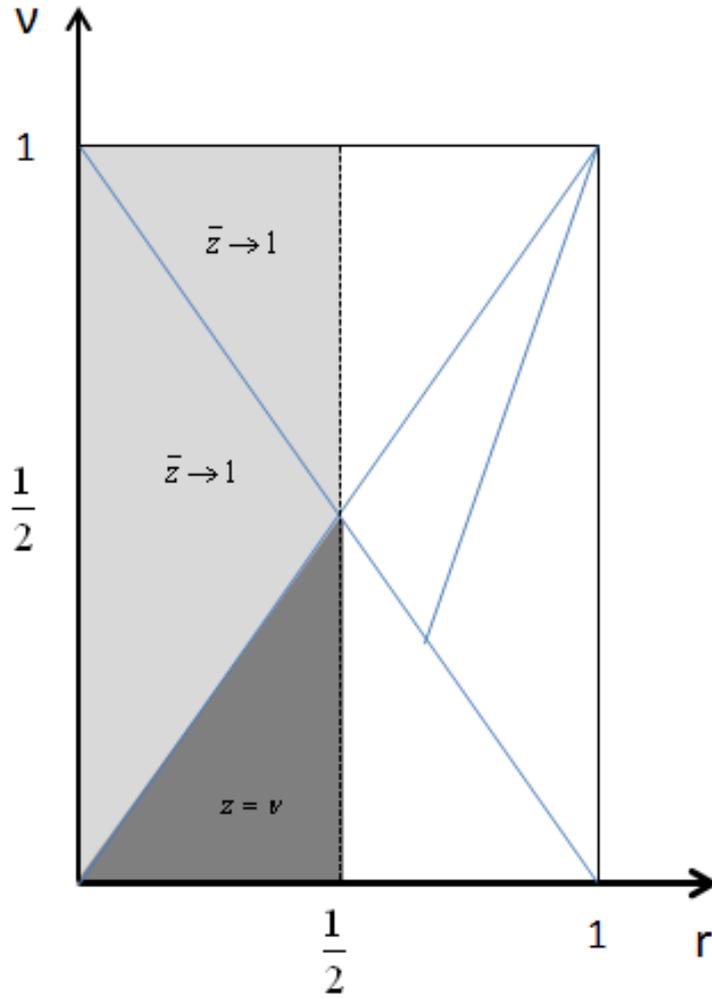


Figure 8: Asymptotic behaviour of the population of payers as a function of the initial condition $z(0) = \nu$ and the parameter $r \leq \frac{1}{2}$ for a WS network with $\hat{k} = N$ for the mixed rule in the uncertainty interval (see main text). Note that the behaviour is symmetric with respect to the value $r = \frac{1}{2}$. If compared with the previous Fig.6, the asymptotic behaviour of this model is simpler since the uncertainty region disappears.

for some time interval, by the (16), whose solution with the initial condition $z(0) = \nu$ is:

$$z(t) = 1 - (1 - \nu) e^{-\nu t}.$$

Therefore, $z(t)$ is always increasing. In consequence, it will cross the value $1 - r$ at some time T and thus, leave the uncertainty region. From this instant on, the ODE that describes the time evolution is:

$$\frac{dz}{dt}(t) = 1 - z(t)$$

whose solution is $z(t) = 1 + \alpha e^{-t}$, where α can be found imposing that $z(T) = 1 - r$. In any case, $z(t)$ will eventually tend to 1. Note that precisely, at $t = T$ the derivative $\frac{dz}{dt}(t)$ jumps from $r\nu$ to r .

Comparing this result with that obtained for the discontinuous rule, in the "mixed case" we have simply that for $\nu < r$, $z(t)$ is constantly equal to ν and for $\nu > r$ then, $z(t)$ tends to 1. As before, a symmetrical situation occurs for $r > \frac{1}{2}$. In conclusion, the pure threshold rule exhibits a richer behaviour than the mixed rule, at least in the mean field approximation.

References

- [1] Andrei, A.L., K. Comer and M. Koehler. An agent-based model of network effects on tax compliance and evasion. *Journal of Economic Psychology*. vol. 40, issue C, 119-133 (2014)
- [2] B. Barzel and A.L. Barabási. Universality in network. *Nature Physics* 9, 673-681 (2013) doi:10.1038/nphys2741dynamics.
- [3] Barash, V. The dynamics of social contagion. Doctoral Thesis (Faculty of the Graduate School of Cornell University) (2011)
- [4] Bloomquist K.M. A comparison of agent-based models of income tax-evasion. *Social Science Computer Review*. Vol. 24 no. 4 411-425 (2006)
- [5] Caldarelli, G. Scale free networks. *Complex webs in nature and technology*. Oxford University Press (2007)
- [6] Campbell, E. and M. Salathé. Complex social contagion makes networks more vulnerable to disease outbreaks. *Scientif reports*, 3:1905 (2013)

- [7] Centola, D. and M. Macy. Complex Contagions and the Weakness of Long Ties. *American Journal of Sociology*, Vol. 113, No. 3, pp. 702-734 (November 2007)
- [8] D. Centola. The spread of behavior in an online social network. *Science*, 329, 1194 (2010)
- [9] Dmitry S. Novikov, Els Fieremans, Jens H. Jensen and Joseph A. Helpert. Random walk with barriers. *Nat Phys.* 7(6): 508-514 (2011) doi:10.1038/nphys1936
- [10] Jackson M.E. and D. López-Pintado. Diffusion and contagion in networks with heterogeneous agents and homophily. *Network Science / Volume 1 / Issue 01 / April 2013*, pp 49 - 67 DOI: 10.1017/nws.2012.7
- [11] Falk, A. and U. Fischbacher. "Crime" in the lab-detecting social interaction. *European Economic Review* 46 (2002) 859 - 869
- [12] Granovetter, M.: Threshold models of collective behavior. *The American Journal of Sociology* (6), 1420-1443 (May 1978) DOI 10.2307/2778111
- [13] Gintis, H., S. Bowles, R. Boyd and E. Fehr. Explaining altruistic behavior in humans. *Evolution and Human Behavior* 24, 153-172 (2003)
- [14] Peter E. Kloeden, Christian Pötzsche (Eds.) *Nonautonomous Dynamical Systems in the Life Sciences*. Lecture Notes in Mathematics. Springer International Publishing, Switzerland (2013)
- [15] Kuhlman, C.J., V.S. Anil Kumar, M V. Marathe, S.S. Ravi and D. J. Rosenkrantz. Finding Critical Nodes for Inhibiting Diffusion of Complex Contagions in Social Networks. J.L. Balcázar et al. (Eds.): *ECML PKDD 2010, Part II, LNAI 6322*, pp. 111-127, Springer-Verlag (2010)
- [16] Hagberg, A.A., D.A. Schult and P.J. Swart. Exploring network structure, dynamics, and function using Networkx. *Proceedings of the 7th Python in Science Conference (SciPy)* (2008)
- [17] Hallam, T.G. and C.E. Clark. Non-autonomous logistic equations as models of populations in a deteriorating environment. *Journal of Theoretical Biology*, Volume 93, Issue 2, 303-311 (1981)

- [18] Hodas, N.O. and K. Lerman. The simple rules of social contagion. *Scientific Reports*, 4, 4343 DOI: doi:10.1038/srep04343 (2014)
- [19] Llacer, T., F.J. Miguel, J.A. Noguera and E. Tapia. An agent-based model of tax compliance: an application to the Spanish case. *Advs. Complex Syst.* 16, 1350007 DOI: 10.1142/S0219525913500070 (2013)
- [20] Lloréns, M. J. C. Nuño, Y. Rodríguez, E. Meléndez-Hevia and F. Montero. Generalization of the Theory of Transition Times in Metabolic Pathways: A Geometrical Approach. *Biophysical Journal* Volume 77, 23-36 (1999)
- [21] Pastor-Satorras, R. and A. Vespignani. Immunization of complex networks. *Phys. Rev. E* 65, 036104 (2002).
- [22] Perc, M., K. Donnay and D. Helbing. Understanding Recurrent Crime as System-Immanent Collective Behavior. *PLoS ONE* 8(10): e76063. doi:10.1371/ journal.pone.0076063 (2013)
- [23] Pickhardt, M. and A. Prinz. Behavioral dynamics of tax evasion û A survey. *Journal of Economic Psychology* 40, 1û19 (2014)
- [24] Shakarian P., Sean Eyre and Damon Paulo. A Scalable Heuristic for Viral Marketing Under the Tipping Model. arXiv:1309.2963v1 [cs.SI] 11 Sep 2013
- [25] Schelling, T.C. *Micromotives and Macrobehavior*. W.W. Norton and Co. (1978)
- [26] Schweitzer F, Mach R. The Epidemics of Donations: Logistic Growth and Power-Laws. *PLoS ONE* 3(1): e1458 (2008) doi:10.1371/ journal.pone.0001458
- [27] Valente, T. W. Network interventions. *Science*. 337, 49 (2012). DOI: 10.1126/science.1217330
- [28] Vespignani, A. Modelling dynamical processes in complex socio-technical systems. *Nat. Phys.*, VOL 8, JANUARY 2012.
- [29] Watts, D.J. and S.H. Strogatz. Collective dynamics of small-world networks. *Nature* 393, 440 (1998).

- [30] Watts, D.J. Six degrees: the science of a connected age. Random House: London, UK (2003)
- [31] Zaklan G., F.W.S. Limab and Frank Westerhoff. Controlling tax evasion fluctuations. *Physica A* 387, 5857-5861 (2008)
- [32] Zaklan, G., F. Westerhoff and D. Stauffer. Analysing tax evasion dynamics via de Ising model. *Journal of Economic of Coordination and Interaction*, 4, 1-14 (2009)