



An agent-based model of network effects on tax compliance and evasion



Amanda L. Andrei¹, Kevin Comer^{*}, Matthew Koehler²

The MITRE Corporation, 7515 Colshire Drive, Mclean, VA 22102-7593, USA

ARTICLE INFO

Article history:

Available online 18 January 2013

JEL classification:

C63
H26

PsycINFO classification:

3020
3040

Keywords:

Social networks
Social influence

ABSTRACT

Agent-based models are flexible analytical tools suitable for exploring and understanding complex systems such as tax compliance and evasion. The agent-based model created in this research builds upon two other agent-based models of tax evasion, the [Korobow, Johnson, and Axtell \(2007\)](#) and [Hokamp and Pickhardt \(2010\)](#) models. The model utilizes their rules for taxpayer behavior and apprehension of tax evaders in order to test the effects of network topologies in the propagation of evasive behavior. Findings include that network structures have a significant impact on the dynamics of tax compliance, demonstrating that taxpayers are more likely to declare all their income in networks with higher levels of centrality across the agents, especially when faced with large penalties proportional to their incomes. These results suggest that network structures should be chosen selectively when modeling tax compliance, as different topologies yield different results. Additionally, this research analyzed the special case of a power law distribution and found that targeting highly interconnected individuals resulted in a lower mean gross tax rate than targeting disconnected individuals, due to the penalties inflating the mean gross tax rate in the latter case.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Neoclassical mathematical models of tax behavior conclude that to maximize their incomes, taxpayers will avoid declaring their actual incomes ([Allingham & Sandmo, 1972](#); [Yitzhaki, 1974](#)), a result that overpredicts what is observed in the real world. This is due to some of the underlying assumptions of the neoclassical models, such as perfectly rational actors and infinite computing capacity ([Axtell, 2007](#); [Kirman, 1992](#)). As an alternative, agent-based models provide more flexibility for analyzing complex systems and collective behavior arising from individual interactions. This research focuses on building an agent-based model in order to examine the impact of social network structures on aggregate tax compliance so that future models may incorporate appropriate networks, thereby resulting in more accurate estimates of individual and collective taxpaying behavior ([Albin & Foley, 1992](#); [Epstein, 2006](#); [Axtell, 2000](#)).

Section 2 provides a background on the problem of tax evasion, complexity theory, agent-based models of tax evasion, and social networks. The next section describes the environmental features, agent characteristics, and rules of the model. Within the results section, the dynamics between different networks are discussed, as well as a special “Big Fish” case, which

^{*} Corresponding author. Tel.: +1 703 983 9898; fax: +1 703 983 1758.

E-mail addresses: aandrei@mitre.org (A.L. Andrei), ktcomer@mitre.org (K. Comer), mkoehler@mitre.org (M. Koehler).

¹ Tel.: +1 703 983 6847.

² Tel.: +1 703 983 1214.

focuses on the impact of power law networks on taxpayers' compliance. The discussion section includes commentary on findings, broader implications, and potential work for the future.

It should be noted that the authors are not presenting a realistic taxing regime that is then evaluated, as such "actual" rates of tax evasion are not provided for comparison. The authors do note, however, that the results of the model are plausible given a 2008 United States Internal Revenue Service estimate of an 84 percent voluntary tax compliance (US Department of the Treasury, 2009). Rather, the authors present a highly stylized taxing regime in order to highlight one main feature: the effect of changing the information flow among taxpaying entities. This being the case, readers should bear in mind several particularly strong assumptions contained in the model, namely, (1) a perfectly flat tax rate and (2) a penalty function used when taxpayers are caught not paying taxes, which can grow without bounds as a taxpayer's income increases.

2. Background

2.1. The problem of tax evasion

Andreoni, Erard, and Feinstein (1998) provide a comprehensive review of literature on tax compliance, most of which focuses on a taxpayer who chooses to declare income, and the reactions of tax authorities and law enforcement to the taxpayers' reports. There also exists research on non-filers, such as the finding of Erard and Ho (2001) that non-filers often hold occupations which make income that is less visible to tax agencies.

Psychological factors such as notions of guilt and shame (Erard & Feinstein, 1994), tax morale (Frey & Torgler, 2007; Alm & Torgler, 2006), social factors such as knowledge of successful evasion (Vogel, 1974), social norms (Alm, Sanchez, & de Juan, 1995), and business ethics (Molero & Pujol, 2012) may also influence taxpayer decisions to comply with tax laws. Additional information on the economic psychology aspects of tax behavior is compiled by Kirchler (2007). This work suggests that the movement of information among a set of social agents is critically important to tax compliance, as well as economic decisions in general (see generally: Easley & Kleinberg, 2010; Jackson, 2008). The underlying network, or how the social agents are connected, therefore is important as it has an impact on the way information can propagate (Dodds & Watts, 2004; Centola, 2010).

2.2. Complexity theory and agent-based models

For the purposes of this research the system of taxpaying behavior is treated as a complex adaptive system, a perspective selected for the following reasons. First, the taxpaying system is comprised of heterogeneous actors such as taxpayers, tax preparers, and tax enforcers. Moreover, each individual within these broad categories is unique, maintaining different values for income, tax rates, risk aversion, etc. The idea of a representative agent in this context is not meaningful (Epstein, 2006). Second, the actors change their behaviors over time. The actors are boundedly rational (Simon, 1991), leading agents to act rationally on the basis of their perception of their environment, rather than according to the objectively best response. Third, the system displays near-decomposability, meaning that although the system is made up of subcomponents, their individual behavior in isolation does not represent the behavior of the overall system when fully interconnected. Finally, as a consequence of the aforementioned characteristics, the system displays emergence ((Crutchfield, 1994)). Although many definitions of emergence exist, for the purposes of this work a definition consistent with Holland (1995) is used: emergence is taken to mean that the behavior of the system is difficult to infer from the behavior of individual components in isolation.

The above characteristics make closed form analysis difficult; therefore, simulation was chosen as an approach to reach a quantitative understanding of this phenomenon. Specifically, the analytic technique of agent-based modeling was chosen (Epstein & Axtell, 1996; Axtell, 2000; Epstein, 2006). Agent-based models are typically made up of three basic components: agents, interaction rules, and space (this could be geo-space or some other abstract space) (Cioffi-Revilla, 2010; Epstein, 2006). As the simulation progresses, agents interact with each other, update their internal states, and may interact with their environment. This creates a coupling among the agents that produces an aggregated dynamic from the micro-level interactions (Axtell, 2005). Given the heterogeneity of the system, adaptation of the agents, agents that would commonly be classified as outliers and excluded from analysis may actually drive the system to particular states otherwise not realized, which may be highly important from a policy standpoint (Schelling, 1978). Within this framework, modelers have interpreted tax compliance factors into a variety of agent characteristics and behaviors, functions and heuristics, and virtual landscapes and networks (see *infra*).

2.3. Agent-based models of tax evasion

Several key models paved the way for computational social scientists and researchers to study tax compliance. The model of Mittone and Patelli (2000) examines how different initial instantiations of heterogeneous types of agents cause variation on collective tax evasion. The model defines three types of agents, each with a unique utility function which defines them as honest, imitative, or free riding. The authors find that the absence of audits causes aggregate non-compliance even among initially honest taxpayers, as public goods begin to diminish and taxpayers withdraw their support for those resources.

Additional experimentation with types of agents demonstrates that varying amounts of public goods become available based on additional revenue raised.

In contrast, [Davis, Hecht, and Perkins \(2003\)](#) initially categorize their model's population of agents as honest or evading. Although the authors do not specifically define the agents' social networks, agents become susceptible to tax evading behaviors if they notice their neighbors profiting from evasion. Evading agents transform into honest agents until they notice their neighbors successfully evading again. The model is used to determine if there exists a critical state, where a specific audit rate changes a population from majority compliant to evasive, or vice versa. Although the model demonstrates that full compliance is reached at audit rates as low as 0.03, the authors admit that they are unable to identify the optimal audit rate which can also be validated by real world data.

Instead of creating a taxonomy of agents, the Tax Compliance Simulator developed by [Bloomquist \(2006\)](#); [Bloomquist, 2008](#), tests hypotheses of tax payment under various law enforcement regimes. Parameters tested include the taxpaying population's changes and reactions to apprehension rates, penalty rates, income visibility, auditor efficacy, and enforcement celerity. Findings included that audit-based deterrence is influenced by social networks: the larger the social network of an agent (i.e., the more neighbors an agent has), the greater the compliance rate of the society.

More notable models recently created include the Networked Agent-Based Compliance Model (NACSM) by [Korobow et al. \(2007\)](#) and [Hokamp and Pickhardt model \(2010\)](#). The NACSM instantiates a single type of agent who may choose to report all, underreport, or not report any income. The model explores the relative impact of a simple social network on taxpaying behavior and tests the influence of individual and collective behavior on a taxpayer. The model also experiments with agent reactions to audits, apprehension, penalties, and fines. [Korobow et al.](#) conclude that a society converges to compliance when taxpayers focus on their own individual decisions and pay little attention to their neighbors. However, in the presence of social networks, the population remains largely non-compliant.

The [Hokamp and Pickhardt model](#) depicts four types of taxpayers: maximizers, imitators, ethical filers, and confused filers. These agents are endowed with an exponential utility function, thereby allowing the model to make more realistic assumptions about audit probabilities than allowed by the traditional utility function. The additional feature of a time lapse allows for more realistic results as well. Results of the model suggest that a time lapse with regards to apprehension, when apprehended agents must account for multiple years of underreporting, is one of the most effective tools for tax compliance.

Finally, it is worth recognizing the models within the domain of econophysics. [Zaklan, Westerhoff, and Lima \(2008, 2009\)](#); [Lima and Zaklan \(2008\)](#); and [Lima \(2010\)](#) utilize the Ising model of ferromagnetism, which in physics describes the interaction of particles when different temperatures are applied. In applying these models to the study of taxpaying behavior, agents do not have individual characteristics or a utility function. Instead, agents can exist in one of two possible states, as compliant or evasive. The influence of neighbors on a taxpayer changes the "social temperature", which decreases or increases the probability of the stochastic "spin-flip" of a taxpayer's state. These models have found that law enforcement has significant influence in directing a population towards tax compliance, even at low levels and despite strong group influence ([Zaklan et al., 2009](#)) and that tax evasion differs among social networks ([Lima & Zaklan, 2008](#)). In particular, the [Zaklan model](#) extended by [Lima \(2010\)](#) was found to be robust for different network structures, with the recommendation that Barabási-Albert power law networks were most effective in simulating tax evasion in the [Zaklan model](#).

2.4. Social networks

Networks are defined as a set of items composed of vertices, also known as nodes, and connections between them, also known as edges or links ([Newman, 2003](#)). These nodes can represent many things, including: people, places, or objects, from individuals to institutions, cities to landmarks, or even particles and artifacts. The edges between these nodes can represent the relationships (or lack thereof) between objects, the flow of information or ideas, or influence between nodes. These edges can be assigned weights and probabilities to enhance the complexity or realism of the model of interest. Edges can also be uni-directional or bi-directional, meaning that there exists a one- or two-sided relationship or influence between nodes ([Schwartz, Cohen, Avraham, Barabási, & Havlin, 2002](#)).

There are multiple metrics for measuring and analyzing networks. One of the most important concepts within network theory is that of centrality, the property of a network which addresses which nodes are the most central and critical ([Newman, 2010](#)). In particular, this research focuses on the ideas of betweenness centrality and closeness centrality. Betweenness centrality measures the number of shortest paths between all nodes that pass through a given node, while closeness centrality refers to the average distance from a node to all other nodes to which it is connected ([Wasserman & Faust, 1994](#)).

As another useful framework within complexity theory for analyzing behavior and interactions, network theory is widely applied throughout the physical and social sciences, and social network analysis has been successful in explaining relationships and interactions between the individual, organizations, and society ([Scott & Carrington, 2011](#)). Social network analysis has been used to examine friendships ([Moreno, 1934](#); [Rapoport & Horvath, 1961](#)), business communities ([Moreno, 1934](#); [Galaskiewicz, 1985](#)), and labor markets ([Granovetter, 1974](#); [Montgomery, 1991](#)), as well as many other types of cultural and socioeconomic connections.

These techniques have also been used to analyze unethical behavior in organizations ([Brass, Butterfield, & Skaggs, 1998](#)), a concept useful for the shadow economy of tax evaders. The structure of the relevant taxpayer network is poorly understood, and researchers speculate on the social structure of taxpayers, often using more simple network topologies in their models. For example, many agent-based models instantiate agents in ring worlds or lattice structures, where an agent has a certain

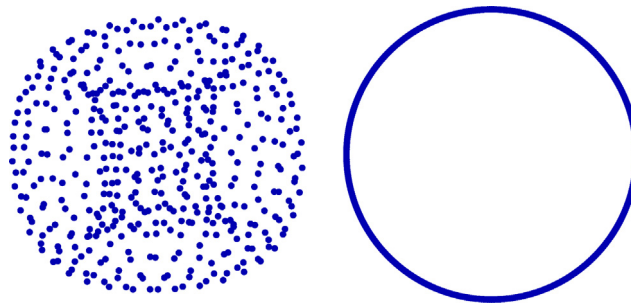
radius of neighbors. Before additional work was expended to determine the correct social network for taxpayers, the authors found it necessary to demonstrate that tax compliance was sensitive to network topology. This paper tests the effects of different network structures, which are defined below and visualized in Fig. 1.

No network consists of isolated nodes, i.e., there are no connections between entities within the set space.

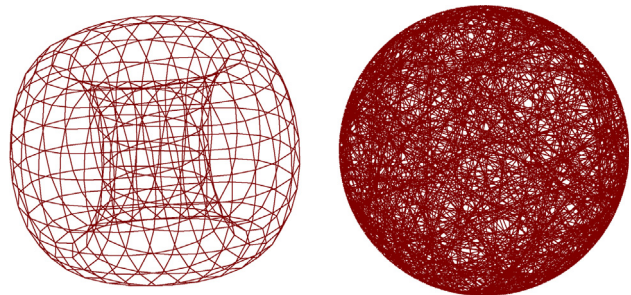
Von Neumann neighborhoods are a common structure in two-dimensional cellular automata models. An agent with a von Neumann neighborhood has four neighbors in the cardinal directions: one to the north, east, south, and west, creating a diamond-shaped pattern on a graph (Weisstein, n.d.b).

Moore neighborhoods are another common structure in two-dimensional cellular automata models. An agent with a Moore neighborhood has eight neighbors in all of the cardinal and ordinal directions, forming a square-shaped pattern on a graph (Weisstein, n.d.b).

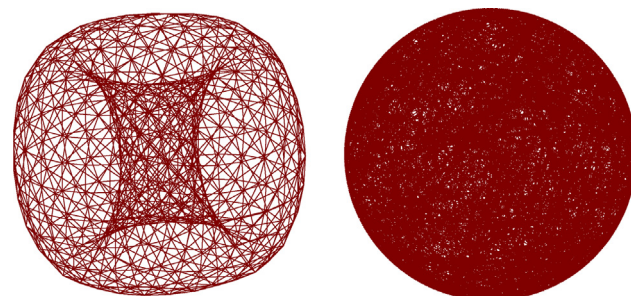
Ring world networks are closed networks which are comprised of nodes that are connected to one node on either horizontal side (Boccaletti, Latora, & Moreno, 2006), i.e., one connected node on the east, and one connected node on the west. When this one-dimensional structure is mapped onto a two-dimensional toroidal surface, the agent structure can be visualized as a helix bent around a circle to close in on itself.



(a) No Network

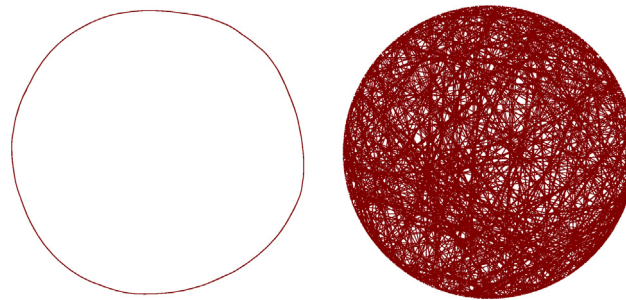


(b) Von Neumann Network

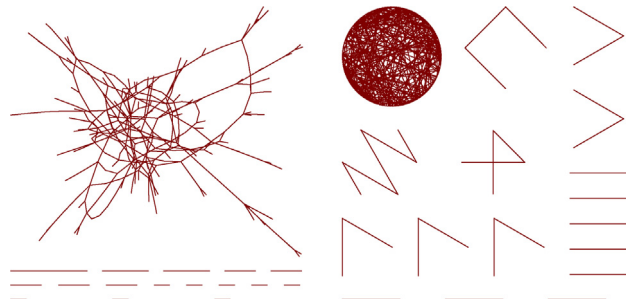


(c) Moore Network

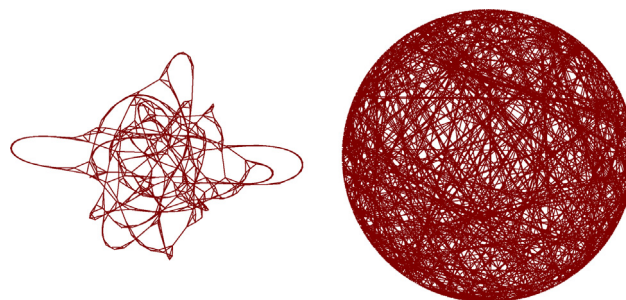
Fig. 1. Different network structures: Each subfigure is a network depicted in two different ways. On the left, a Spring Embedding node layout algorithm is used. On the right, the Circular Embedding algorithm is used. These two views were created to give a cleaner view of the networks (on the left) and a consistent view of the networks (on the right). The line segments are disconnected groups of agents in pairs or triplets. Images created by the authors with NetLogo and Mathematica. (a) No network. (b) Von Neumann network. (c) Moore network. (d) Ring network. (e) Erdős–Rényi network. (f) Small Worlds network. and (g) Power Law network.



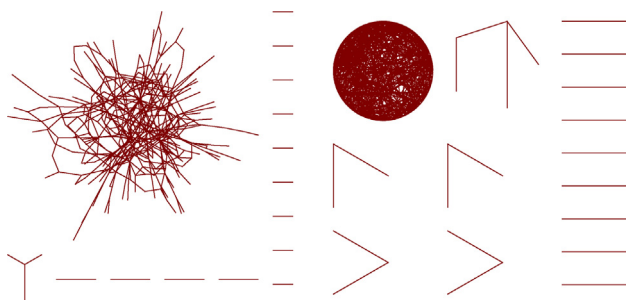
(d) Ring Network



(e) Erdős-Rényi Network



(f) Small Worlds Network



(g) Power Law Network

Fig. 1 (continued)

Erdős-Rényi networks consist of vertices that, barring multiple collections, are connected randomly. Extensions of this network have included connecting vertices with certain probabilities, or including a non-Poisson degree distribution (Boccaletti et al., 2006). They are also known as random graphs.

Small Worlds networks are usually generated on low-dimensional lattices. A fraction of links between nodes are broken and rewired with some probability to another node (Newman, 2003). Characterizing the world as “small” asserts that

agents are somehow “close” to each other. These networks are numerically large, decentralized, and highly cohesive and clustered (Watts, 1999). They are also known as Watts–Strogatz networks.

Power Law networks have a power law distribution of edges per node (Clauset, Shalizi, & Newman, 2009), where most nodes have only a few connections, but a small fraction of nodes are highly and disproportionately connected. No single node can represent this network. The network can remain stable if random nodes are removed, but the network is susceptible to directed attacks on highly connected nodes (Andriani & McKelvey, 2007). These distributions are also known as scale-free, as the shape of the distribution does not change across many orders of magnitude.

As a general note, von Neumann and Moore neighborhoods are useful for modeling spatial or geographic concepts. While the other networks noted above can be utilized this way, they are often used to construct space in an abstract way, e.g., depicting a friendship network with long edges between nodes, thereby denoting levels of distance between friends.

Additionally, the authors chose Clauset et al. (2009) power law networks rather than Barabási and Albert (1999) networks in order to highlight potential structural differences in the networks. If Barabási–Albert networks were used, then there would be no isolates. As noted earlier, this study focuses on the impact of changes to information flow among taxpaying entities. The authors chose to use the most “extreme” various networks. As tax evasion and penalization involve covert behavior and punishment, it seemed desirable to the authors to test network structures that were not fully connected.

Much evidence has been found that networks are important to taxpaying behavior. Furthermore, much work has been done showing that the flow of information changes with network structure. This study is an effort to explicitly connect these two lines of research by exploring the impact of network structure on a system of stylized taxpaying behavior.

3. Model description

3.1. Hypothesis

The null hypothesis (H_0) for this research states that networks cause no change in the Mean Gross Tax Rate (MGTR). That is, across the seven types of network structures tested, the mean of the population of each network μ_x is such that

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7$$

The alternative hypothesis (H_a) is that the MGTR differs across populations using different networks, or

$$H_a : \mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4 \neq \mu_5 \neq \mu_6 \neq \mu_7$$

The Mean Gross Tax Rate will serve as the primary metric for tracking the effects of networks across the simulation runs. This term is derived from the Voluntary Mean Tax Rate (VMTR), provided in Hokamp and Pickhardt (2010).

As this model includes revenue from taxes, as well as penalties collected from apprehended agents, into the declared income for each agent, “voluntary” is hardly an appropriately descriptive term for this model. Therefore, the metric has been termed Mean Gross Tax Rate (MGTR), to represent the revenue received from both taxes and penalties. The equations for the MGTR are discussed in more detail in Section 4.1.

It should be noted that this is not the method by which penalties are collected by the Internal Revenue Service in the United States. Taxes and penalties are not collected or even measured in tandem as a single entity in the real world. The model aggregates these two factors to note the differences across the two networks, both in agent behavior and in aggregate revenue generation.

3.2. Environment features

The model in this research was implemented using NetLogo, a highly intuitive software toolkit and programming language used to build agent-based models (Wilensky, 1999). This modeling tool can instantiate large numbers of agents with complex behaviors and display their environment in an easily manipulated graphical user interface. Parameter sweeps are conducted through “BehaviorSpace”, another tool of NetLogo to allow for systematic sensitivity analysis.

This model depicts the dispersion and effects of tax evasion across a variety of social network topologies. Several of its features such as agent types, time lapses, and rules for declaring and apprehension are derived from the Hokamp and Pickhardt (2010) and Korobow et al. (2007) models which are elaborated upon in Section 3.3. This section focuses on the spatial and temporal features of the model.

Initially, agents are randomly distributed in a two-dimensional toroidal surface without any social structure. One of seven types of networks can be applied, including that of no network. The authors used the definitions of networks as defined previously, with some specifications for the following cases. In the Ringworld network, agents made connections not only to the nearest node to the east and west, but also to the next node beyond the nearest node. This creates a system, where each agent continued to have four neighbors, similar to the von Neumann network, in order to create a network that more closely emulated the network used in the Hokamp model. For the Erdős–Rényi network, the number of edges varies between 100 and 1000. For the Small Worlds network, the probability of connectivity varies between 0% and 100%. Additionally, it is the

only network that used a probability in generating the links between nodes. That is, to make a Small Worlds network, a Ring-world network was constructed and then reconnected at 10% of the edges, chosen randomly.

There is also an element of time in this model. The model simulates 40 years of tax payments and apprehension. For the purposes of this model, those 40 years are considered to be a full “tax cycle”. As agents are apprehended, the memory of being apprehended remains with them for a number of years, and they are unlikely to underreport in the near future. However, as the memory of the apprehension fades over time and the agent is influenced by other agents, the agent may again underreport.

3.3. Agent features and rules

For the reasons outlined above, the agents in the model are adaptive, bounded rational, and embedded in social networks. For performance reasons and to fit within the NetLogo graphical interface, the agent population consists of exactly 441 agents, fifty of which are “honest” and always declare their actual incomes, and fifty of which are “dishonest” and calculate the lowest possible incomes they can declare as based on their risk aversions and their subjective probabilities of apprehension. The remaining agents are characterized as “imitating” agents, who observe the behavior of agents to whom they are connected and calculate their income based on decisions of their neighbors.

As adaptive agents, these taxpayers react and change their decisions of how to file their taxes based on events that affect them directly, such as being apprehended. Apprehension refers to the discovery and subsequent penalization by enforcement agencies of agents who declare less than their actual income. No false positives are assumed in this situation; all apprehended agents are guilty of declaring less than their actual income. If an agent is apprehended, it will increase its subjective apprehension probability while the objective apprehension rate operates externally. This objective apprehension rate represents the rates set by government and law enforcement institutions and is completely unknown to the agents. When an agent is apprehended, it pays a hefty penalty which is proportional to its actual income.

Agents make their decisions of how to file based on the parameters in Table 1 and the following equations. These equations are derived from the Hokamp and Pickhardt and Korobow et al. models and adapted to the purposes of this model; namely, to simplify a complex tax regime and demonstrate the effects of social networks.

Honest agents always declare their actual incomes. Therefore an agent declares income according to

$$X_{i,t} = W_{i,t} \quad (1)$$

A dishonest agent calculates the lowest possible income they can declare as based on its risk aversion and its subjective probability of apprehension. The following two equations are from Hokamp and Pickhardt model (2010).

If

$$\rho_s < \frac{\theta}{(\theta + (\pi - \theta)e^{\lambda\pi W_{i,t}})} \quad (2)$$

then $X = 0$.

If

$$\rho_s > \frac{\theta}{\pi} \quad (3)$$

then $X_{i,t} = W_{i,t}$, as individual risk is too great, and the dishonest agent will declare their actual income.

However, if ρ_s exists between the values as given by (2) and (3), the declared income becomes

$$X_{i,t} = W_{i,t} - \frac{\ln([(1 - \rho_s)\theta]/[\rho_s(-\theta + \pi)])}{\lambda\pi} \quad (4)$$

Imitating agents observe the behavior of agents to whom they are connected. They calculate their income based on the product of the average of their neighbors' ratio of actual to declared income and their own declared income, such that

Table 1
Model parameters and descriptions.

Parameter	Description	Value
θ	Tax rate	0.30
π	Penalty rate	0.45
X	Declared income	Defined
W	Actual income	$\sim U(0, 100)$
ρ_s	Subjective probability of an apprehension	Dynamic per Markov process
ρ_o	Objective probability of an apprehension	0.02 for graphs, tested {0, 0.02, 0.04, 0.06, 0.08, 0.1}
λ	Individual risk aversion	$\sim U(0.00, 1.00)$
v	Number of neighbors linked to each agent	Depends on network

$$X_{i,t} = \frac{1}{v} \sum_{j=i-v}^{i-1} \frac{X_{j,t-1}}{W_{j,t-1}} W_{i,t} \quad (5)$$

If an agent is not apprehended, and its $\rho_{s,t} > \rho_o$, then its subjective apprehension probability decreases according to

$$\rho_{s,t+1} = \rho_{s,t} - 0.2 \quad (6)$$

This equation is based on the human decision-making heuristic of availability (Tversky & Kahneman, 1973) and the fact that humans are averse to loss (Kahneman & Tversky, 1984). We assume that as time goes on and agents are not apprehended, the agents' perception of being apprehended decreases because an agent assumes itself to be safer as time passes without any apprehension to itself or its neighbors. It should be noted that agents never know the objective probability, and therefore do not compare the subjective probability to the objective probability. Additionally, the code restricts the subjective probability to between 0 and 1: as soon as an agent's subjective probability is less than zero, it is reset to the objective probability.

If an agent is apprehended, it adjusts its declared income and subjective apprehension probability to the following:

$$X_{t+1} = \theta(W_t - X_t)(1 + \pi W_t), \quad \rho_{s,t} = 1 \quad (7)$$

4. Results

4.1. Dynamics between different networks

Two metrics were collected over the various runs of the model: MGTR and the number of agents apprehended.

MGTR is derived from Hokamp and Pickhardt's definitions and equations for VMTR, which is computed from the average amount of each agent's Voluntary Tax Rate (VTR), as computed by the equation

$$VTR = \theta \frac{X_i}{W_i} \quad (8)$$

One should note that the X_i term in Eq. (8) includes penalties; therefore, *ceteris paribus*, a more honest population may actually create a lower MGTR as fewer penalties will be collected.

Therefore, MGTR, the average value over all agents at a given time, is expressed as the following for n agents:

$$MGTR = \theta \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n W_i} \quad (9)$$

The second metric, agents apprehended, refers to agents that were caught declaring an income lower than their actual income.

Each parameter set consisted of 25 runs. The parameter set used one of the seven networks defined in Section 2.4, as well as a specific apprehension rate. The values of apprehension rates are defined in Table 2. Across these different rates and networks, there were a total of 42 different parameter sets. Each run consisted of 40 steps which represents 40 years, i.e., a full tax cycle. Metrics were collected over all 40 steps in order to observe the dynamics of a complete tax cycle. The system was also scaled to test a population of 50 agents and 5000 agents, and there was no statistically significant change in the differences observed in network behaviors.

With the exception of "No Network", at a 2% apprehension rate, the dynamics of agent behavior show an agent population that may initially be less compliant, but over time converges to a steady state, as demonstrated in Fig. 2. Agents with no network display little change in behavior over the course of the simulation, primarily because agents have no influence on each other. With no network, imitating agents default to reporting all of their income.

The dynamics for the von Neumann, Moore, Ringworld, and Small Worlds networks are relatively similar to each other. As agents start at a low MGTR, they are eventually apprehended and pay penalties for evading taxes. These enforcement activities propagate via the network and create feedback, increasing the amount of declared income within the population as a whole. By step 10, the population has reached an equilibrium with little fluctuation in the level of tax compliance.

The Power Law and Erdős–Rényi network dynamics differ slightly from the other networks tested. While the simulation begins with the same activation phase and rate of apprehension, these network populations do not reach a steady state as rapidly as the other network populations. Additionally, the steady state to which they converge is of a higher value than the other networks, resulting in a MGTR that is higher than the actual tax rate. Although these results appear counter-intuitive, they can be explained by the impact of penalties upon the agents. In this model, penalties incurred for evading taxes can be quite steep. For example, given $\theta = 0.3$, $\pi = 0.45$, $W = 100$ (in units of the US median annual income of about 50,000 dollars) and $X_t = 0$ in Eq. (7), the penalty can be $X_{t+1} = 1380$ after being apprehended, which is almost fourteen times the agent's actual income. As the example demonstrated, depending on how large the difference is between the agent's actual income and the agent's declared income, the monetary punishments can potentially amount to multiple times the actual income of the agent.

Table 2
Parameters tested and corresponding values.

Parameter	Value
Type of network	No network von Neumann Moore Ringworld Erdős–Rényi Small Worlds Power Law
Rate of apprehension	{0, 0.02, 0.04, 0.06, 0.08, 0.1}

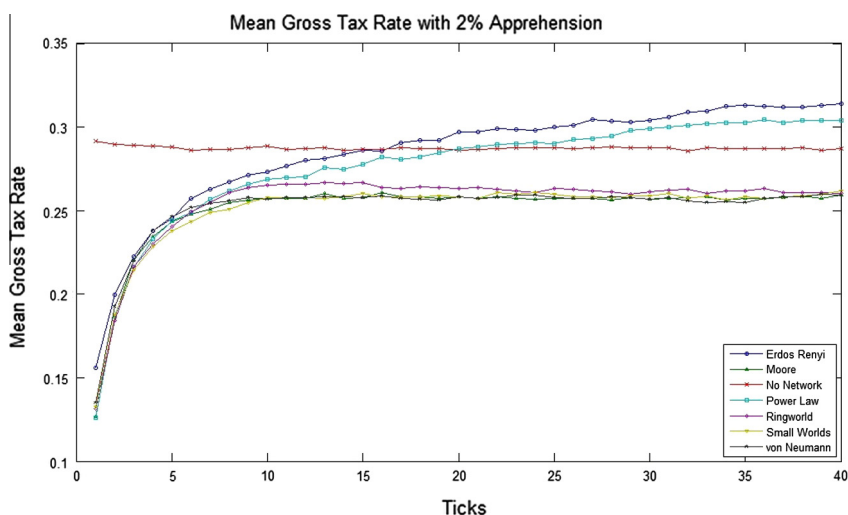


Fig. 2. Comparison of mean gross tax rate over time across different networks, at 2% apprehension rate, $\theta = 30\%$.

There are no statistically significant differences in how many agents are apprehended depending on which network they belong. The number of taxpayers apprehended appears to not be an influential factor on aggregate tax compliance. Rather, agents evade taxes to a much higher degree in certain networks, therefore causing more agents to be apprehended, which in turn leads to significantly higher penalties being collected over the agent set.

The effects of the various networks connecting the agents is exacerbated when the apprehension rate increases. In Fig. 3, the activation dynamics of the simulation is similar to the dynamics observed at the 2% apprehension rate, with agents being apprehended, paying a penalty, and returning to equilibrium. Meanwhile, agents in the Power Law and Erdős–Rényi networks eventually reach an equilibrium, but at a much higher mean gross tax rate than other networks.

Fig. 4 displays the variance, standard deviations, median, and outliers of the MGTR across the different network topologies. As demonstrated, Power Law and Erdős–Rényi graphs have the greatest variance about their means. For analysis, this study examined the primary null hypothesis presented earlier in this paper. A single-variable ANOVA test produced an F -value of 131 and a $\text{Prob} > F$ of 1.099×10^{-60} . Therefore, assuming a 95% confidence level, one is able to reject the null hypothesis that there is no statistical difference in the means of the networks.

A secondary test was implemented, examining the difference in means across paired networks. For networks a and b , the null hypothesis held that there is no statistical difference between the means of networks a and b , or $H_0: \Delta_{a,b} = \mu_a - \mu_b = 0$. The alternative hypothesis is that the difference of the means is statistically significant, $H_a = \Delta_{a,b} \neq 0$. This was tested using a two-tailed, single-variable ANOVA test. The 95% confidence intervals for the difference of the true means of various pairs of networks are displayed in Table 3.

If a confidence interval contains the value of 0, then the study failed to reliably reject the null hypothesis, and the networks are statistically similar. On the other hand, if the confidence interval does not contain 0, then the study probably rejects the null hypothesis, and the difference of the means of the two networks is statistically significant.

4.2. The “Big Fish” case

In order to test the impact of closeness of the network on the MGTR via agent apprehension, the Power Law network was used for an additional set of runs and analyses. The BigFish runs, as the name implies, targets those “bigger” individuals that have many more connections in the Power Law network than other nodes and are often referred to as the hubs of a social

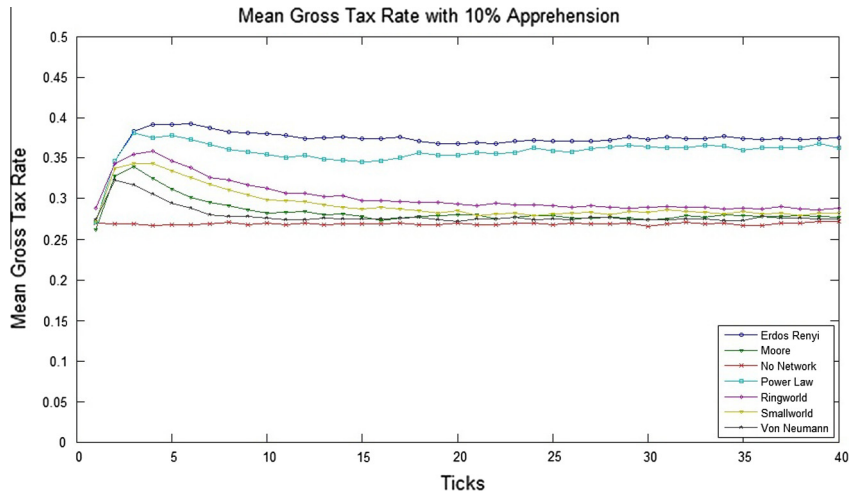


Fig. 3. Comparison of mean gross tax rate over time across different networks, at 10% apprehension rate, $\theta = 30\%$.

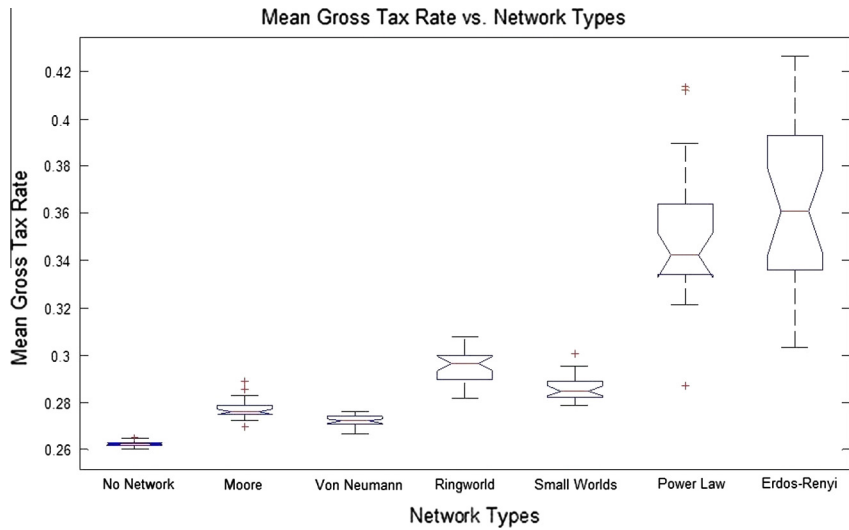


Fig. 4. Displays the variances, standard deviations, medians, and outliers for MGTR vs. Network Types at 10% apprehension rate. The tops and bottoms of each box are the 25th and 75th percentiles of the samples, respectively. The distances between the tops and bottoms are the interquartile ranges. The line in the middle of each box is the sample median. The whiskers are lines extending above and below each box. Whiskers are drawn from the ends of the interquartile ranges to the furthest observations within the whisker length (the adjacent values). Observations beyond the whisker length are marked as outliers (more than 1.5 times the interquartile range away from the top or bottom of the box), and are displayed with a red + sign (MATLAB documentation).

network. In this model, usually eight to ten agents have the top five highest values for number of links, and are selected as “big fish”. These highly connected agents are constantly monitored and apprehended with 100% probability when they declare less than their actual income.

The BigFish runs were compared against non-BigFish runs (see Tables 4 and 5), using the same apprehension schema, but selecting them at random rather than by the quantity of links that an agent owns. The null hypothesis stands that there is no statistical difference between the results of these two runs, as defined by the MGTR; $H_0: \Delta = \mu_1 - \mu_2 = 0$. The alternative hypothesis is that the BigFish runs will produce a statistically significantly different MGTR than the non-BigFish runs; $H_1: \Delta = \mu_1 - \mu_2 \neq 0$.

Given the results shown in Table 4, there is enough evidence to reject the null hypothesis at the 95% confidence level, and to state that apprehending big fish creates a statistically different result in the tax evasion behavior of the Power Law network, the statistical characteristics which are shown in Table 5.

The point of this experiment was to prove that the networks connecting heterogeneous agents are an important component to understanding tax evasion dynamics. The authors do not assume the conclusions are quantitatively correct, given the large scale of the penalties incurred by some agents. That aside, it is not unreasonable to consider penalties to be truly monstrous for certain cases of tax evasion, such as repeat offenses, or tax evasion paired with other financial crimes.

Table 3

95% confidence intervals for the difference of the true means of various pairs of networks. (*) Denotes statistical significance.

	No network	Moore	von Neumann	Ringworld	Small Worlds	Power Law	Erdős–Rényi
No network		{0.0004, 0.0292} (*)	{-0.0045, 0.0243}	{0.0182, 0.0469} (*)	{0.0091, 0.0378} (*)	{0.0722, 0.1009} (*)	{0.0865, 0.1153} (*)
Moore			{-0.0193, 0.0095}	{0.0034, 0.0321} (*)	{-0.0057, 0.0230}	{0.0574, 0.0861} (*)	{0.0717, 0.1005} (*)
von Neumann				{0.0083, 0.0370} (*)	{-0.0008, 0.0279}	{0.0623, 0.0910} (*)	{0.0767, 0.1054} (*)
Ringworld					{-0.0234, 0.0053}	{0.0396, 0.0684} (*)	{0.0540, 0.0827} (*)
Small Worlds						{0.0487, 0.0774} (*)	{0.0631, 0.0918} (*)
Power Law							{0.0000, 0.0287} (*)
Erdős–Rényi							

Table 4

Results of the two-tailed t-test between BigFish and non-BigFish simulations at a 95% confidence level.

Statistic	Value
$P(\Delta = 0)$	1.6434×10^{-9}
95% Confidence interval	{0.0058, 0.0107}
t-Statistic	6.6474
Degrees of freedom	99
Sample standard deviation	0.0123

Table 5

Comparison of BigFish and non-BigFish runs.

Parameter	BigFish	Non-BigFish
Maximum	0.2838	0.3179
Mean	0.2658	0.2741
Minimum	0.2469	0.2571
Median	0.2663	0.2722
Standard deviation	0.0073	0.0109
Variance	0.0001	0.0001

5. Discussion

5.1. Findings

The primary focus of this experimental study was to characterize the sensitivity of tax evasion models to the social network topology of the agents contained within them. The results clearly demonstrate that networks play a considerable role in the collective behavior of the agent population.

The values of Power Law and Erdős–Rényi networks demonstrate a statistically significant difference when compared to the other networks. However, the differences between the results from the Power Law and the Erdős–Rényi networks are not statistically significant. The other networks among themselves demonstrate a statistically significant difference. Therefore, the type of network structure implemented is critical when modeling tax evasion.

It is worth asking: why do Power Law and Erdős–Rényi networks display such different results compared to the other networks and arrive at such vastly different results? This can primarily be explained by the centrality of the Power Law and Erdős–Rényi networks and the assumptions of the model. One type of centrality, closeness centrality, refers to the sum of the shortest distance from one node to all other nodes. This type of centrality allows agents to imitate each other in this model. Due to the specific nature of these networks, information and imitative behavior can spread in fewer time steps to widely dispersed agents than through other types of network structures tested here. In the case of this model, agents are rapidly sharing information about the amount they declare with respect to how much they actually own. Signaling behavior, such as an agent acting honest after apprehension, is also spread much more rapidly through the Power Law and Erdős–Rényi networks than other networks.

Another form of centrality is betweenness centrality, which measures the number of shortest paths that connect certain nodes. Fig. 5 displays the relationships between these networks based upon the aforementioned centrality measures, and Fig. 6 compares the number of connections between the different network structures. Because agents do not “gossip” in our model, the impact of high betweenness is limited. This is not the case for Power Law and Erdős–Rényi networks. In these topologies, there are a number of highly connected hubs which allows for influences to propagate much further than what

might normally be expected in other networks. The high closeness centrality observed in the Power Law networks and the Erdős–Rényi networks exacerbates the impact of very honest or very dishonest agents. In these high closeness centrality networks, the “hubs” have an impact on a large number of agents. If the hubs are honest, the population will tend to be more honest. On the other hand, if the hubs are dishonest, the population will tend towards dishonesty.

5.2. Broader implications

Throughout this paper the differences seen among the various network structures were highlighted. There is a potential practical application that can be learned from these results. If tax authorities target more highly interconnected individuals, then one should expect the “honesty” of the population to increase (as least within the stylized system depicted in this simulation study). In order to test this hypothesis, two sets of specialized runs were performed, the BigFish study, discussed *supra*. For these runs only the power law network was used, as it contained highly interconnected hubs, disconnected singletons, and bilateral networks. In one set of runs the tax authority audited the hubs every year for the 40 year run. In the other set of runs the tax authority audited a number of disconnected nodes that was equal to the number of hubs in the network (typically eight to ten agents).

If the above hypothesis is correct, there should be a *lower* MGTR when the tax authority audits highly interconnected agents than disconnected agents. This is due to the fact that the hub agents are connected to many agents, who then observe the hub agents declaring all of their income. Therefore, the imitating agents that are connected to them will be more likely to declare more of their income. Based on how MGTR was calculated for this research, this results in a *lower* MGTR than observed when auditing disconnected agents. Recall in this study, penalties collected by the tax authority was included in the MGTR calculation. This being the case, a more honest population would generate fewer penalties and, therefore, likely produce a lower MGTR than would a less honest population.

This is exactly the dynamic that was observed in these runs (see Table 5). The null hypothesis (p -value of 1.6434×10^{-9}) that there is no difference between the tax authority enforcement schemes can be firmly rejected. As anticipated, the MGTR for the hub auditing scheme is lower than that of the disconnected agent auditing scheme, indicating that the overall population is more honest in the hub auditing scheme. Moreover, hubs may be “created” by a tax authority publicizing successful audits.

Additionally, with regards to studying tax evasion or economic agent-based models in general, an important implication of this study concerns the examination and use of the proper network topology to connect agents. Any description of an agent-based model, or another model utilizing network theory, should define what network topology is being used, and why that structure is best suited to address the issue at hand.

5.3. Future work

In order for this research to have effective contributions to the reduction of tax evasion, the dynamic between enforcement and tax evasion must be established. As enforcement enacts new policies to penalize tax evaders, agents adapt by redistributing their income across various accounts, asking for financial information from among their social network, or

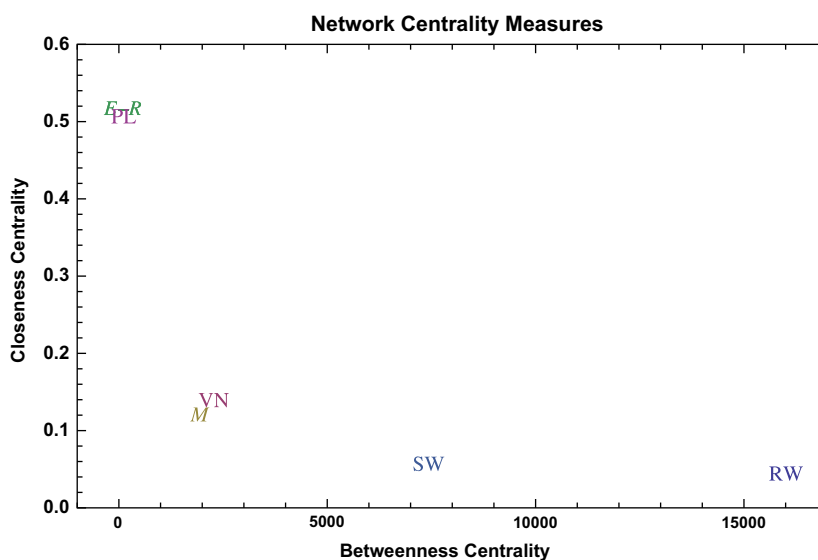


Fig. 5. Scatter plot of network centrality measures (betweenness vs. closeness). ER = Erdős–Rényi, PL = Power Law, VN = von Neumann, M = Moore, SW = Small Worlds, RW = Ringworld.

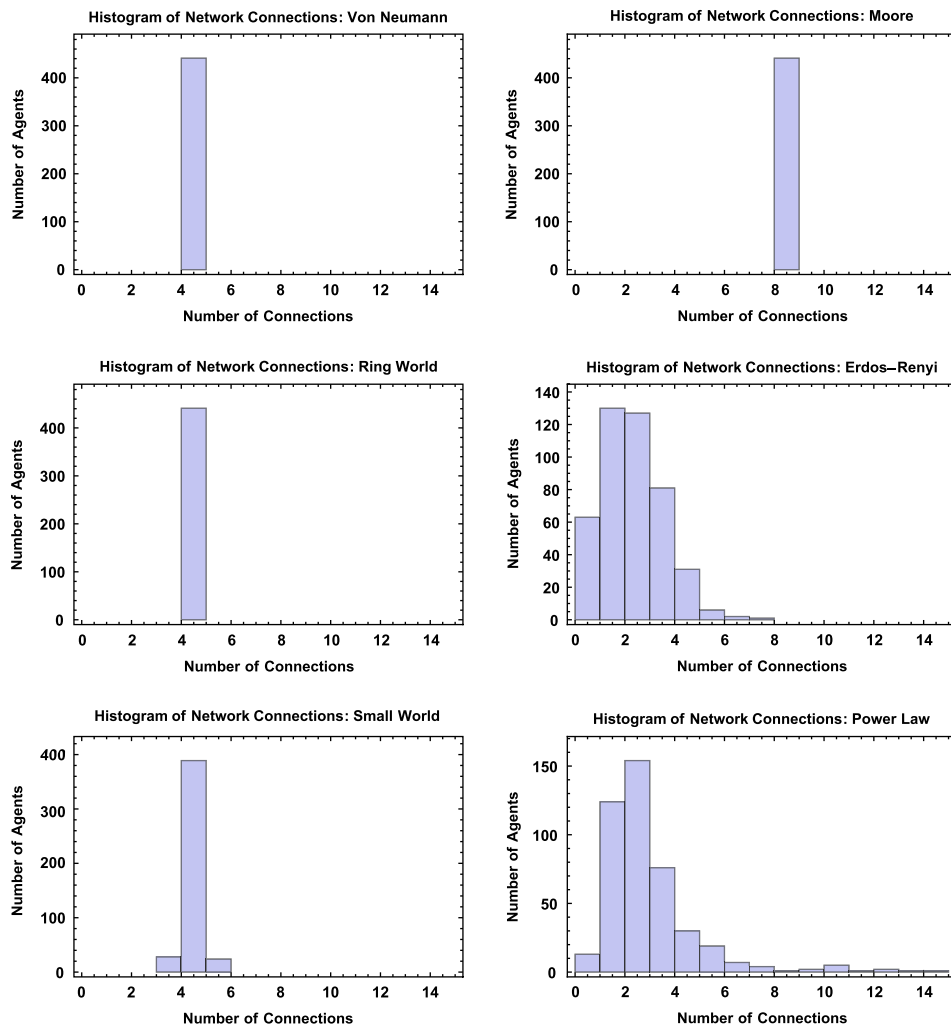


Fig. 6. Comparison of histograms of distribution of connections of different networks.

becoming honest in their income declarations. From these agent adaptations emerges a complex system of tax evaders and tax enforcement reacting to the actions of the other to achieve certain goals. Therefore these dynamics need to be explored further, using data to validate the behaviors implemented in any study.

The work described here demonstrates the importance of information within the tax paying population and the structure of the network over which the information flows. However, this work only looked at the dynamics of enforcement and did not address the dynamics of tax evasion. The importance of the coevolution between enforcement and evasion should not be understated. The adaptation between enforcers and evaders within this system is what created the complex dynamics and will be the subject of ongoing work.

Furthermore, the networks included in this study, while well documented in graph theory, were not based on any data collected on real world social networks of taxpayers. In this case, the application of networks was purely theoretical. If the data collected in various tax studies was incorporated into this model, regarding both the penalties associated with tax evasion and the relevant social network among taxpayers, the model could offer additional concrete insight into the propagation of tax evading behavior among various taxpayers.

6. Concluding remarks

This research has built upon two important agent-based models of tax evasion, the NACSM developed by [Korobow et al. \(2007\)](#) and the model developed by [Hokamp and Pickhardt \(2010\)](#). The agent-based model created in this research used the apprehension rules of the NACSM and the behavioral rules from [Hokamp and Pickhardt \(2010\)](#) in order to test the impacts and effects of seven types of network structures on aggregate tax compliance. It was discovered that there are in fact two

network structures that significantly influence taxpaying behavior: the Erdős–Rényi network and the Power Law distributed network. In these network structures, information and influence is propagated and disseminated much more quickly than in the five other structures tested. This is attributed to the closeness centrality of the networks, which allows information about maximum payoffs to travel through shorter distances and to many more agents at once.

This model was highly stylized and had many of the nuances of real world tax regimes removed in order to demonstrate the impact social networks upon tax compliance and evasion. While some of the equations are more abstract than what might be found in more realistic models of evasion, this is a necessary step in exploring the complexity of the tax system in more depth and detail. Having completed this step, we plan to refine this model by including assumptions and equations that accurately reflect real world tax systems as embedded in their own socioeconomic and political contexts.

These findings have profound implications for understanding tax evasion. Future models of tax compliance should explicitly state and explain what network structures they implement, as well as why those topologies were chosen. Furthermore, policymakers may have better insight into taxpaying behavior if they know the potential social structures of individuals, communities, organizations, and institutions. While it is outside the scope of this research to provide policy recommendations, these methodologies and findings provide a foundation for future exploration of tax compliance in a myriad of disciplines and fields.

References

- Albin, P., & Foley, D. (1992). Decentralized, dispersed exchange without an auctioneer. *Journal of Economic Behavior and Organization*, 18(1), 27–51.
- Allingham, M. G., & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1, 323–338.
- Alm, J., Sanchez, I., & de Juan, A. (1995). Economic and noneconomic factors in tax compliance. *Kyklos*, 48, 3–18.
- Alm, J., & Torgler, B. (2006). Culture differences and tax morale in the United States and Europe. *Journal of Economic Psychology*, 1(3–4), 323–338.
- Andreoni, J., Erard, B., & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature*, 36(2), 818–860.
- Andriani, P., & McKelvey, B. (2007). Beyond Gaussian averages: Redirecting organization science toward extreme events and power laws. *Journal of International Business Studies*, 38, 1212–1230.
- Axtell, R. (2000). Why agents? In *Proceedings of the agents 2000 conference*.
- Axtell, R. (2005). *Three distinct kinds of empirically-relevant agent-based models*. Working Paper, Brookings Institution Center on Social and Economic Dynamics.
- Axtell, R. (2007). What economic agents do: How cognition and interaction lead to emergence and complexity. *Review of Austrian Econ*(20), 105–122.
- Barabási, A., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
- Bloomquist, K. (2008). Taxpayer compliance simulation: A multi-agent based approach. In B. Edmonds, C. Hernandez, & K. Troitzsch (Eds.), *Social simulation: Technologies, advances and new discoveries*. Hershey, PA: IGI Global.
- Bloomquist, K. M. (2006). A comparison of agent-based models of income tax evasion. *Social Science Computer Review*, 24(4), 411–425.
- Boccaletti, S., Latora, V., Moreno, Y., et al (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424, 175–308.
- Brass, D. J., Butterfield, K. D., & Skaggs, B. C. (1998). Relationships and unethical behavior: A social network perspective. *Academy of Management*, 23(1), 14–31.
- Centola, D. (2010). The spread of behavior in an online social network experiment. *Science* (329), 1194–1197.
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, 51, 661–703.
- Cioffi-Revilla, C. (2010). Computational social science. *Computational Statistics*, 2(3), 259–271.
- Crutchfield, J. P. (1994). The calculi of emergence: Computation, dynamics, and induction. *Physica D, Proceedings of the Oji International Seminar Complex Systems – From Complex Dynamics to Artificial Reality*, 75, 11–54.
- Davis, J. S., Hecht, G., & Perkins, J. D. (2003). Social behaviors, enforcement and tax compliance dynamics. *Accounting Review*, 78, 39–69.
- Dodds, P., & Watts, D. (2004). Universal behavior in a generalized model of contagion. *Physical Review Letters*, 92(21).
- Easley, D., & Kleinberg, J. (2010). *Networks, crowds, and markets*. Cambridge, MA: Cambridge University Press.
- Epstein, J. (2006). *Generative social science*. Princeton, NJ: Princeton University Press.
- Epstein, J., & Axtell, R. (1996). *Growing artificial societies*. Washington, DC: Brookings Institution Press.
- Erard, B., & Feinstein, J. (1994). Honesty and evasion in the tax compliance game. *RAND Journal of Economics*, 25(1), 1–19.
- Erard, B., & Ho, C. (2001). Searching for ghosts: Who are the non-filers and how much do they owe? *Journal of Public Economics*, 81, 25–50.
- Frey, B., & Torgler, B. (2007). Tax morale and conditional cooperation. *Journal of Comparative Economics*, 35(1), 136–159.
- Galaskiewicz, J. (1985). *Social organization of an urban grants economy*. New York, NY: Academic Press.
- Granovetter, M. (1974). *Getting a job: A study of contacts and careers*. Cambridge, MA: Harvard University Press.
- Hokamp, S., & Pickhardt, M. (2010). Income tax evasion in a society of heterogeneous agents – Evidence from an agent-based model. *International Economic Journal*, 24(4), 541–553.
- Holland, J. H. (1995). *Hidden order*. New York, NY: Helix Books.
- Jackson, M. (2008). *Social and economic networks*. Princeton, NJ: Princeton University Press.
- Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist*, 39(4), 341–350.
- Kirchler, E. (2007). *The economic psychology of tax behavior*. New York, NY: Cambridge University Press.
- Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives*, 6(2), 117–136.
- Korobow, A., Johnson, C., & Axtell, R. (2007). An agent-based model of tax compliance with social networks. *National Tax Journal*, 60(3), 589–610.
- Lima, F. (2010). Analysing and controlling the tax evasion dynamics via majority-vote model. *Journal of Physics: Conference Series*, 246.
- Lima, F., & Zaklan, G. (2008). A multi-agent-based approach to tax morale. *International Journal of Modern Physics C: Computational Physics & Physical Computation*, 19(12), 1797–1808.
- Mittone, L., & Patelli, P. (2000). Imitative behavior in tax evasion. In F. Luna & B. Stefansson (Eds.), *Economic simulations in swam: Agent-based modeling and object oriented programming* (pp. 133–158). Amsterdam: Kluwer Academic Publishers.
- Molero, J., & Pujol, F. (2012). Walking inside the potential tax evader's mind: Tax morale does matter. *Journal of Business Ethics*, 105(2), 151–162.
- Montgomery, J. (1991). Social networks and labor-market outcomes: Toward an economic analysis. *American Economic Review*, 81(5), 1408–1418.
- Moreno, J. L. (1934). *Who shall survive?* Beacon, NY: Beacon House.
- Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45, 167–256.
- Newman, M. E. J. (2010). *Networks: An introduction*. New York, NY: Oxford University Press.
- Rapoport, A., & Horvath, W. J. (1961). A study of a large sociogram. *Behavioral Sciences*, 6, 279–291.
- Schelling, T. (1978). *Micromotives and macrobehavior*. New York, NY: W.W. Norton & Company, Inc.
- Schwartz, N., Cohen, R., Avraham, D. ben, Barabási, A.-L., & Havlin, S. (2002). Percolation in directed scale-free networks. *Physical Review E*, 66, 015104.
- Scott, J., & Carrington, P. (2011). *The SAGE handbook of social network analysis*. London: Sage Publications Ltd.
- Simon, H. (1991). Bounded rationality and organizational learning. *Organization Science*, 2(1).

- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2), 207–232.
- US Department of the Treasury. (2009). Update on reducing the federal tax gap and improving voluntary compliance. Washington D.C.
- Vogel, J. (1974). Taxation and public opinion Sweden: An interpretation of recent survey data. *National Tax Journal*, 27, 499–513.
- Wasserman, S., & Faust, K. (1994). *Social network analysis*. Cambridge: Cambridge University Press.
- Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 105(2), 493–527.
- Weisstein, E. W. (n.d.b) Moore neighborhood. From MathWorld — A Wolfram Web Resource. <<http://mathworld.wolfram.com/MooreNeighborhood.html>>.
- Weisstein, E. W. (n.d.b). Von Neumann neighborhood. From MathWorld — A Wolfram Web Resource, 16. <<http://mathworld.wolfram.com/vonNeumannNeighborhood.html>>.
- Wilensky, U. (1999). *Netlogo*. Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. <<http://ccl.northwestern.edu/netlogo/>>.
- Yitzhaki, S. (1974). A note on income tax evasion: A theoretical analysis. *Journal of Public Economics*, 3, 201–202.
- Zaklan, G., Westerhoff, F., & Lima, F. W. S. (2008). Controlling tax evasion fluctuations. *Physica A: Statistical Mechanics and its Applications*, 387(23), 5857–5861.
- Zaklan, G., Westerhoff, F., & Stauffer, D. (2009). Analysing tax evasion dynamics via the Ising model. *Journal of Economic Interaction and Coordination*, 4, 1–14.