Tax enforcement in an agent-based model with endogenous audits

Susanna Calimani and Paolo Pellizzari

Abstract We generalize the classic Allingham and Sandmo's model of tax evasion considering heterogeneous agents with different degrees of tax morale and matchable, as opposed to non-matchable, income. The Tax Agency evolves its control scheme, maximizing the revenues from fines, and takes into account some minimal information on the taxpayers. We compare different audit policies and find that the most effective scheme remarkably depends on the way agents update the subjective probability of being audited, on the distribution of matchable income in the population as well as on the level of tax morale. Hence, different features of societies and taxpayers' behaviors not only affect the compliance rate, as expected, but require the Tax Agency to alter its audit policy in a context-dependent way. In particular, high revenues are obtained performing random audits when agents think they are directed towards peculiar individuals and, conversely, should be biased towards low declarations when taxpayers believe audits are nonspecific or random.

1 Introduction

Tax evasion has always been a dear issue to policy makers, but in times of crisis when Governments fall short of resources, it easily becomes one of the favorite pieces in everybody's political programme. It is clear that tax evasion dynamics should be investigated taking into account the joint action of many heterogeneous taxpayers and of the Government (or, equivalently, of the agency that collects taxes). However, starting with the seminal paper [Allingham and Sandmo, 1972], taxpayers' behavior has been at the centre of the focus, using models with fully rational

Susanna Calimani

Dept. of Economics and Statistics Cognetti de Martiis, University of Turin. e-mail: susanna.calimani@unito.it

Paolo Pellizzari

Dept. of Economics , Ca' Foscari University of Venice. e-mail: paolop@unive.it

agents who decide how much of their actual income to declare to the tax authority, in an expected utility framework *à la* von Neumann and Morgenstern. The conduct of the Government was instead simply mimicked using a constant fine rate and assuming that entirely random auditing is performed with some probability. This basic model well captures the deterrent effects of increased audit probability and fines, but typically predicts much higher evasion than it is observed and suggests that hikes in the tax rate would increase compliance, whereas intuition would lead to the opposite outcome. An impressive number of alternative models tried to overcome these shortcomings, resorting, say, to the tendency of *overweighting small probabilities*, [Kahneman and Tversky, 1979], introducing a psychological cost of evasion ("shame") in the utility function and developing more general paradigms in which the *tax morale* of a community plays a role, see [Torgler, 2002]. The interaction between agents and Government can also be framed in terms of power and trust, see [Kirchler et al., 2008], where the former increases the *enforced* compliance and the latter can boost *voluntary* compliance.

In this paper, we exploit the flexibility and heterogeneity of an agent-based approach to show that the inclusion of more realistic features notably improves the descriptive accuracy of the standard model. The agent-based model enables us to depict a system of many heterogeneous agents, whose interaction might give rise to emergent phenomena that are not analytically derivable, in particular when the model exhibits potential network effects, nonlinear behavior, learning or adaptation. These features characterize a multitude of social phenomena, and certainly also the tax evasion framework we examine: in such a contest, we use agent-based simulations to grasp the effects that different combinations of societal configurations and audit policies might have.

First, we allow the tax agency to evolve its own control scheme, using minimal information about taxpayers' income to improve the efficacy of the audits. The weight of these pieces of information can, in fact, be calibrated to optimize who is to be inspected.

Second, our agents are restricted in their evasion decisions and, in particular, we assume that some fraction of income cannot be concealed. Income can be categorized as traceable or non-traceable: the first includes salary or wage as well as self-employed income that can be matched by somebody else's tax-report, whereas the second is formed by all those earnings' components that are hardly matchable and easily hidden. Therefore, we emphasize in this work the distinction of matchable versus non-matchable income, as done in [Bloomquist, 2003] where it is empirically shown that the non-matchable component notably increased in recent years for US taxpayers, with visible repercussions on aggregate compliance. Take, for instance, two taxpayers with the same gross income: the first earns 90% of his revenues from matchable wage, whereas the second gets 90% of income from non-matchable sources. Clearly, it is materially possible for the latter to strategically conceal a vast amount of her income and devise profitable tax evasion. Fresh evidence from a study conducted in Denmark has indeed corroborated Francis Bacon's saying that *opportunity makes a thief*: tax evasion is found to be substantially higher for individuals who self-report their income compared to those who have most part

of their resources reported by a third party to the tax agency, regardless of audits probabilities, [Kleven et al., 2011].

Third, we allow population's risk aversion to be correlated with income and with the portion of matchable wealth. This allows for the customization of the audit policy to different states of the economy, as in booming phases the correlation of risk aversion and income in the whole population is likely to be negative, whereas the link may be weaker in gloomy periods.

Fourth, we test two ways to sense and adapt the perceived probability of being audited. Taxpayers can either estimate this probability using a sample average based on random matches with other peers or using their own history of past audits. We call these schemes *geographical* and *temporal* adaptation, respectively. It turns out that the efficacy of audit schemes critically depends on how agents subjectively estimate their audit probability in the next period.

On the one hand, taxpayers may, in this setup, evade less than theoretically predicted just because they cannot do otherwise. On the other hand, this model fully incorporates in the game a tax agency who is able to use some (minimal) pieces of information about the agents to maximize the revenues from its audit policy: this looks realistic, as more sophisticated schemes than purely random inspections are clearly within reach for a tax agency. Other works investigated some simple endogenous audit schemes [Collins and Plumlee, 1991, Alm et al., 1993], finding significantly lower levels of underreporting, but letting the score depend on the degree of matchable income appears, to the best of our knowledge, novel and promising.

This paper is organized as follows. Section 2 provides details on the ways taxpayers and the tax agency are modelled. The following section discusses our simulation results and show how alternative audit policies interact with the distribution of matchable income, tax morale and agents' beliefs regarding the probability of verification. Section 4 provides some conclusive remarks and relates our findings with the existing literature, discussing a few policy suggestions.

2 The model

We model a society made of a tax agency (TA) who collects taxes from N individuals. The next two subsections will provide a detailed description of the agents and of the society in which they are embedded. Such a society, as it is often the case in agent-based models, depends on some features of the tax system as well as on the meta-parameters of the distributions individual traits are drawn from.

2.1 Agents

Agents have exogenously given parameters, $I_j, \rho_j, \kappa_j, \beta_j, j = 1, \ldots, N$, held constant across time, which denote privately known income, risk aversion, tax morale and the

fraction of matchable income, respectively. Individual parameters are sampled only once at the beginning of the simulation from a distribution described in the next subsection. Taxpayers decide what portion of their earnings d_{jt} ($0 \leq d_{jt} \leq 1$) to disclose at time $t = 1, ..., T$ in order to maximize their expected utility. The declared income, $d_{jt}I_j$ cannot fall short of their own matchable income $\beta_j I_j$ ($0 \leq \beta_j \leq 1$), which can proxy wage from (legal) employment, third-party reported income or earnings that cannot otherwise be concealed. Agents assume that they will be audited at time t with some probability p_{jt}, whose true value is unknown and must be estimated in one of two ways. The first method, called *geographical*, assumes that each taxpayer meets k "neighbors" per period[1] and learns the number m_t of those who were audited. The previously held $p_{j,t-1}$ is updated as follows:

$$p_{jt} = w p_{j,t-1} + (1-w) \frac{m_{jt} + A_{jt}}{k+1},$$

where $w = 0.5$ and A_{jt} is 1 if the j-th agent was audited at time t and 0 otherwise. In other words, the probability of experiencing an audit is the average of the past estimate, a fresh guess derived from the k encounters and the knowledge of whether she was audited herself. Alternatively, using *temporal* updating, each agent uses a time-average and computes, in any t

$$p_{jt} = \frac{1}{t} \sum_{i=1}^{t} A_{ji}.$$

While geographical updating almost exclusively depends on current audits on other taxpayers, temporal adjustment focuses only on the history of one agent, uses past information and is more accurate if a longer span of time is available.

The utility function of taxpayers is

$$U(d_{jt}) = (1 + d_{jt})^{\kappa_j} W_{jt}^{(1-\rho_j)}$$

where W_{jt} is the wealth after taxes and fines (if any) and κ_j ($0 \leq \kappa_j \leq 1$) represents the tax morale of agent j: the stronger the ethical sense and the more income the agent will report, the higher utility she will perceive. A stronger tax morale corresponds to a marked sensitivity of utility to changes in the choice of how much to conceal: indeed a less *moral* taxpayer, endowed with a lower κ_j, will feel less ashamed in not paying taxes than an upright one.

The fraction d_{jt} of income to be disclosed is optimally selected by each agent solving the problem

$$\max_{d_{jt} \geq \beta_j} EU(d_{jt}) = p_{j,t-1} U(X_{jt}) + (1 - p_{j,t-1}) U(Y_{jt}),$$

---

[1] We independently and uniformly sample the k neighbors from the whole population at each time and for each agent.

where $X_{jt} = (1 - \tau d_{jt})I_j - f\tau(1 - d_{jt})I_j$ is the income if audited and $Y_{jt} = (1 - \tau d_{jt})I_j$ is the income in the absence of an audit. In the previous equation, τ and f denotes the tax and penalty rates, respectively. Both parameters characterize the society in which agents live, which is described in the next subsection.

2.2 Society

We define "society" as the set of arrangements and parameters that describe the tax collection process and the distribution of individual traits previously detailed.

The TA collects taxes from agents setting the tax rate τ they must pay on their income and the fine rate f to be levied on hidden income, when and if evasion is discovered. The TA can audit a fixed number qN ($0 \leq q \leq 1$) of agents, according to some policy. We assume that this is done by assigning a score to each taxpayer and picking qN agents with probability proportional to the score. We consider four audit schemes: random auditing simply gives the same score to everyone in every period; strict cutoff sets the score to 1 for the qN agents whose declared income $d_{jt}I_j$ is the smallest and 0 otherwise; in the (mild) cutoff rule, auditing is performed proportionally to the rank of the declared income; finally, by enhanced auditing we refer to a scoring system that is developed by the TA in the attempt of maximizing the revenues from enforcement.

Using strict cutoff, the TA will audit in each period the qN individuals who declared the least. Agents who report low income are likely to be inspected also under cutoff auditing but there is much more variability with respect to strict cutoff and many more individuals experience one or more audits along time.

Enhanced auditing is based on the score S_{jt}, which is a function of matchable and declared income:

$$S_{jt} = p_1(\beta_j I_j)^{p_2} + (d_{jt}I_j)^{p_3},$$

where p_1, p_2 and p_3 are constants selected by the TA to approximately maximize the revenues from audits.[2]. While random auditing is a special case of the enhanced scoring system (set $p_1 = 1, p_2 = 0, p_3 = 0$ to obtain constant scores), only a part of all possible functions of βI and dI is explored with this parametrization. Clearly, neither the strict nor the mild cutoff can be exactly replicated but the relative size of p_2 as compared to p_3 and the scaling factor p_1 can provide some guidance in singling which users are more likely to be audited using the enhanced scheme.

---

[2] The objective of the TA, the maximization of the sum of the fines imposed in qN audits, is a stochastic function of p_1, p_2, p_3 and the approximate solution depends also on the "givens" of the society. This is to say that different enhanced schemes are likely to be developed in different societies or when agents behave differently. We mimic the TA's search for good triplets of p_1, p_2, p_3 by means of an Evolution Strategies algorithm, which is stopped after 30 functions' evaluations and prematurely halts. Therefore, the process should be interpreted as a somewhat realistic quest by a boundedly rational TA of an audit scheme that is tailored to the society and capable of improving the revenues from fines. For additional details on Evolution Strategies see [Beyer and Schwefel, 2002], for implementation details see [Trautmann et al., 2011]

Name	Symbol	Distribution
Income	I	Lognormal, mean=30000, s. d. ≈ 23500
Tax morale	κ	Uniform in $[\kappa_{low}, 1 - \kappa_{low}]$, $\kappa_{low} = 0, 0.025, ..., 0.25$
Risk aversion	ρ	Uniform in $[0, 1]$
Matchable income	β	beta(a, a), $a = 0.5, 1, 2$

Table 1 Description of the societal individual parameters' distributions.

The description of the society is completed by the definition of the distributions used to draw the individual parameters, see Table 1. We assume population income to follow a log-normal distribution, with mean income being 30000 and standard deviation equal to 23500 (these figures vaguely reflect Italian ones); the risk aversion parameter is uniformly distributed in $[0, 1]$. For the sake of simplicity, we suppose that tax morale is uniformly distributed among the taxpayers and let the support of the density change symmetrically around the mean, thus effectively considering mean preserving "spreads" of the same distribution. This way we can easily represent different societies with distinct moral attitudes: on average tax morale has the same value in every country, but we can find societies with more extreme values -in both directions- than others and increments in κ_{low} increase aggregate compliance. The fraction of matchable income β_j for the j-th agent follows a beta(a, a) distribution and, in particular, we focus on three specific values of $a = 0.5, 1, 2$ describing density functions that are U-shaped, uniform and bell shaped, respectively. When $\beta \sim$ beta$(0.5, 0.5)$ most of the agents have either high or low matchable income, whereas there is low density for middle ways; on the contrary, a beta$(2, 2)$ distribution corresponds to a society where taxpayers mostly have a mix of matchable and non-matchable income, with few extreme cases. A country like Italy, where self-employment often leaves many opportunities for income disguising, can be representative of the first scenario ($a = 0.5$), whereas a nordic country, say Norway, where usually payments are completed by traceable means could better be approximated by the second situation ($a = 2$). The case relative to $a = 1$ stands in between.

Finally, we capture important second order effects in the distribution of citizens's individual traits of one society allowing for nontrivial correlation of parameters. Hence, while Table 1 reports the marginal distributions of parameters, we assume that $Cor(\rho_j, \beta_j) = r$ and $Cor(\rho_j, I_j) = -r$ across the population. Picking, say, $r > 0$ is tantamount to suppose that more risk-averse agents have on average a smaller income. At the same time, they are likely to have a larger matchable income. In most of the countries we can think of, this positive correlation appears to be reasonable to account for the self-selection process that generally leads risk-averse agents to seek employed job and risk-prone taxpayers to self-employ in more profitable activities that also leave more room for evasion.[3]

---

[3] Technically, we obtain correlated marginals as follows: we sample from a 3-dimensional multivariate normal with the given correlations; once we have a normal vector $(z_j^{(I)}, z_j^{(\rho)}, z_j^{(\beta)})$ for each agent, we invert the appropriate cumulative probability distribution (log-normal, uniform and beta, respectively) to obtain (I_j, ρ_j, β_j) with the desired approximate correlations.

Since the TA does not make q public, taxpayers are not fully informed about the true intensity of audits and have a perceived probability of auditing p_{jt} that is updated according to one of the two possible ways explained before. Furthermore, taxpayers do not know how the tax agency actually runs the inspections and which algorithm is followed when selecting the tax files to audit.

3 Results

The model presented in the previous section generalizes the classic framework in several ways: agents are restricted in their compliance decision and must declare at least as much as their matchable income; they can update the perceived probability of audit using geographical or temporal adjustments, while the correlation among parameters and the level of tax morale of the population varies. We simulate a grid of beta distributions, where $a \in \{0.5, 0.75, ..., 2\}$ (7 values), and tax morale levels $\kappa_{low} \in \{0.025, 0.050, ..., 0.250\}$ (11 values). Each grid was then replicated 4 times, to account for two possible values if correlations among parameters, $r = 0$ and $r = 0.5$ (2 values), and the two updating schemes, denominated geographical and temporal (2 elements). For each set of values for the meta-parameters a, κ_{low}, r and one updating method, we simulated 30 periods with $N = 1000$ agents, for a total of $7 \times 11 \times 2 \times 2 = 308$ societies, where $\tau = 27\%$ and $f = 1$ are constant. To avoid any dependence on the initial values of the probability p_{j0} at time 0, we discard the first 29 periods and report the results of the last one, which can be thought of as one fiscal year.

As customarily in agent-based models, the richness of the data is both a curse and a blessing and we especially focus in what follows on the audit policies, contrasting the random, enhanced, strict cutoff and mild cutoff audit systems. Figure 1 depicts the revenues from fines of the four policies, when agents geographically update their own subjective probabilities, $r = 0.5$ and $a = 0.5$.

The solid lines are relative to the gains of strict cutoff, enhanced and mild cutoff scoring, normalized by the revenues of the random scheme. For instance, with minimal tax morale ($\kappa_{low} = 0$), the strict cutoff and enhanced auditing produce revenues that exceed those of the random scheme by over 500% and about 350%, respectively, as can be seen in the left part of the picture. When tax morale is low, it is clear that audits based on the strict cutoff rule are more lucrative on average than any other scheme. Notice that this finding points out that the standard model, where only random auditing is performed and tax morale is identically null, is unstable in that the TA would have a strong incentive to change its auditing method. There is substantial variability in the outcomes, as displayed by the dashed lines showing one s.d. of revenues away from the mean: even though the average value is significantly higher under strict cutoff, there is a moderate and definitely non-null probability that, say, enhanced auditing will raise more fines in a single period (i.e., in a random batch of qN audits).
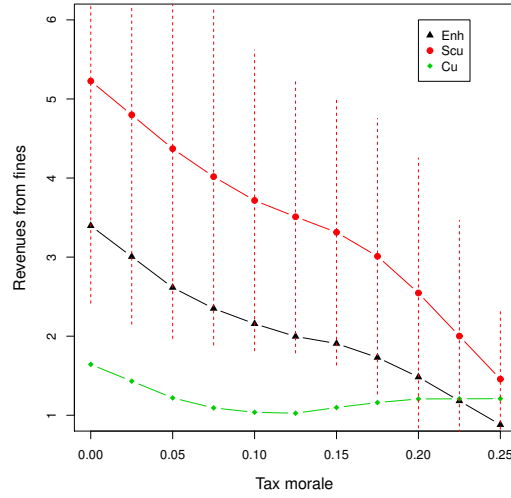
Fig. 1 Revenues obtained by strict cutoff, enhanced and mild cutoff selection methods (from top to bottom, normalized by the revenues of the random audit scheme), relative to geographical updating, $r = 0.5$, $a = 0.5$. The vertical dashed segments depict one standard deviation above and below the mean of revenues for the strict cutoff policy.

In the framework depicted in the figure, despite the effort made by the TA to "maximize" the revenues, the enhanced scheme is not the most profitable. This is due to the high level of noise present in the stochastic objective, to the limited resources allocated for the task (the search stops after 30 evaluations, see Footnote 2) and to specific features of the society.

The high average revenues generated by the strict cutoff rule derive from a somewhat extreme combination of effects: due to positive r, more wealthy taxpayers are likely to have lower risk aversion and are more inclined to evasion; in the cases where their tax morale and matchable income happen to be low, such agents simply conceal virtually all their income and declare $d_j \approx 0$. The strict cutoff rule often samples repeatedly wealthy total evaders and, therefore, revenues are boosted. However, this is possible only when agents geographically update the probability of being audited and "do not realize" that they will probed much more frequently if they declare low incomes. For increasing levels of κ_{low} the difference in performance among the scoring rules markedly decreases, thus showing that revenues from tax enforcement are relatively insensitive to the audit policy in societies with high tax morale (besides obviously being much lower).

Table 2 shows a summary of our results when agents use geographical updating of the probabilities. The data in the table show how the relative revenues (with respect to random auditing) of the three considered policies vary as r, the tax morale, and the distribution of matchable income change in the society. We denote by LTM (HTM) low (high) tax morale societies in which $\kappa_{low} = 0.025$ ($\kappa_{low} = 0.225$).

| | | r = 0.5 | | r = 0 | |
		LTM	HTM	LTM	HTM
	Enh	3.00	1.18	1.53	0.93
a = 0.5	Scu	4.80*	2.00*	4.83*	1.75*
(U-shaped)	Cu	1.43	1.21	1.19	1.00
	Enh	0.96	1.28	1.36	1.09
a = 1	Scu	3.53*	1.78*	3.15*	1.42*
(Uniform)	Cu	1.15	1.54	1.24	1.00
	Enh	0.85	2.61*	1.85	2.28*
a = 2	Scu	1.76*	0.87	2.46*	1.25
(Bell-shaped)	Cu	0.90	1.09	1.55	0.83

Table 2 Tax Agency revenues from fines, when taxpayers *geographically* update p_{jt}. Values are normalized by the revenues of the random audit scheme.

The best performing audit policy (highlighted with *) is quite often the (rather brutal) strict cutoff selecting system. This always holds when tax morale is low, regardless of the distribution of the matchable income and for any value of r. The relative efficacy of the strict cutoff policy decreases moving from a U-shaped ($a = 0.5$) to a bell-shaped distribution ($a = 2$). This result is due to the reduction, as a moves from 0.5 to 2, of the number of "extreme" taxpayers who have, at the same time, low risk aversion, low individual tax morale κ_j and small matchable income. Hence, the strict cutoff rule increasingly audits real poors as opposed to aggressively non-compliant taxpayers.

The inspection of the columns labelled HTM, relative to high levels of tax morale, reveals that the enhanced policy obtains better results in the case of bell-shaped distribution of matchable income. The intuition here is straightforward: high tax morale and few individual with extremely low β requires a more nuanced audit policy, which does not concentrate itself on the smallest declaration but takes a more sophisticated approach in which both $d_j I_j$ and $\beta_j I_j$ have a role in the scoring function. This is confirmed by the values of p_1, p_2 and p_3 that determine the enhanced policy: for increasing a, p_2 and p_3 have the tendency to increase and to cluster around zero, respectively, and this translates into a larger weight given to the βI term as compared to the dI one.[4]

Table 3 presents the simulation results when agents temporally update their subjective assessment of the auditing probabilities. Recall that in this case, taxpayers focus exclusively on their own audits' history, averaging over time the relative frequency of the inspections they have undergone. On the one hand, this disregards the information on the unconditional intensity of control q, which could be estimated by sampling information on the neighbors; on the other hand, however, any given taxpayer is likely to be able to estimate with much greater accuracy the probability of being audited conditional on his behaviour, which is ultimately what she should care about.

The insights that can be obtained by Table 3 remarkably differ from the ones descending from Table 2. We stress that this is only due to the different method

---

[4] We thank an anonymous referee for this remark.

		$r = 0.5$		$r = 0$	
		LTM	HTM	LTM	HTM
	Enh	0.88	1.05	1.64	0.76
$a = 0.5$	Scu	1.20*	0.31	2.02*	0.26
(U-shaped)	Cu	0.83	1.24*	1.44	1.15*
	Enh	0.90	0.92	0.77	0.88
$a = 1$	Scu	1.13*	0.20	0.70	0.24
(Uniform)	Cu	1.09	1.31*	1.18*	1.29*
	Enh	1.53*	1.37*	0.69	0.78
$a = 2$	Scu	0.84	0.27	0.55	0.25
(Bell-shaped)	Cu	1.36	0.74	0.90	0.89

Table 3 Tax Agency revenues from fines, when taxpayers *temporally* update p_{jt}. Values are normalized by the revenues of the random audit scheme.

in updating the subjective audit probabilities. Overall, as several values are smaller than 1, there are numerous cases in which random auditing over-performs at least one competing scoring scheme. When $r = 0$ and $a = 2$, surprisingly, random auditing appears to be the best policy. Generally, the difference in performance across policies is reduced under temporal updating and there is no clearly dominant strategy: strict cutoff is the best option in 3 cases (only), whereas enhanced scoring achieves the best results in 2 cases (for positive r and bell-shaped matchable income). Mild cutoff scoring secures the highest revenues in 5 cases, mostly related to societies with high tax morale. Evidently, as temporal updating allows taxpayer to partially anticipate the decisions of the TA, a good deal of randomness in the choice of the taxpayers to audit has favorable payoffs. Mild cutoff, indeed, can be interpreted as a random scheme endowed with some bias that increases audit probabilities for low declarations; indeed, as noticed before, this audit policy keeps sampling a variety of agents and does not get trapped in repeated audits of the same agents.

A careful scrutiny of the rich set of data available at the micro level shows an interesting dynamics going on between evasion-prone agents and the TA. A taxpayer subject to audit at time t will, under temporal updating, revise upward his belief about the probability $p_{j,t+1}$ of experiencing an audit in the following period. *Ceteris paribus*, this will increase the amount $d_{j,t+1}$ disclosed to the TA at $t + 1$ and reduce the chance of an audit. As a consequence, $p_{j,t+2}$ would decrease on average, pushing the agents to conceal more income and -consequently- increase their likelihood of being audited at $t + 2$ and so on. Such a *hide and seek* game is only possible when temporal updating is used by taxpayers and, in this setup, one of the best options for the TA would be, game theory *docet*, a good deal of randomization with an eye to low levels of *dI*, which is exactly what may be achieved by an enforcement policy guided by the mild cutoff scoring.

4 Conclusions

The agent-based model we have studied in this paper extends the standard model of tax evasion, allowing for heterogeneous taxpayers, consideration of matchable income, the adoption of several alternatives to the common random audit schemes, and two plausible ways to assess and update the subjective probability of being audited. Some (but not all) versions of the previous features were studied in previous works, but our model provides a comprehensive picture of the complex interactions occurring between the TA and the taxpayers. Such a detailed representation is precluded to most analytically solvable models, that have a more limited scope and must be based on simplifying (and often heroic) assumptions.

As a first general remark, the model shows that effective audit policies are dependent on the context. We confirm that the tax morale is an important factor in explaining tax evasion. Indeed, high values of tax morale, nearly always correspond to lower revenues for the TA (due to lower evasion rate). At the same time, the way agents perceive and update the probability of audit is also extremely relevant: if taxpayers take compliance decisions mainly based on audits experienced by others, there is scope for the TA to adopt the lucrative strict cutoff rule that simply audits those who declare the least. On the contrary, a seemingly minor modification in the way probabilities are perceived and updated, that is taking into account the history of each taxpayer, results in a different outcome where most enforcing policies are somewhat similar and mild cutoff, a biased version of random auditing, appears to be the best option to maximize revenues. Realistically, taxpayers will use a mix of the two stylized updating methods we considered but, even in this case, the TA should modify its actions depending on judgements or guesses about both the way agents behave and their tax morale. The temporal updating method is definitely more precise than the geographical one in determining the likelihood of being audited -when this is done not randomly- and, as a matter of fact, when agents become more acute, the TA needs to use more subtle selection methods. As people use geographical updating (and react too little to their own experience), the strict cutoff outperforms almost always the other methods. An exception is the case of beta(2,2) and high levels of tax morale, possibly because agents are willing to declare more (they are more moral) and there are less people with plenty of non-matchable income. The TA needs to be more sophisticated: metaphorically, when you go fishing in a sea with both big and small fish, a wide-mesh net can be used; but if you go to the fishing pond, where fish are not as big and maybe swim deep, then more ingenious means are needed.

Other structural features of the society, such as the distribution of the matchable income or the correlations among individual parameters, have relevant effects. A U-shaped configuration, where many agents have plenty of material opportunity for evasion, should be tackled by targeting low declarations (still keeping robust doses of randomness in the choice). In contrast, bell-shaped distribution of matchable income requires different audit policies and more nuanced approaches, giving more weight to the matchable component and considering the level of tax morale. There

is even one case (see the bottom right panel of Table 3) in which the best policy is simply random auditing.

According to the simulation, the efforts exerted by the TA to develop enhanced auditing rule are, with the minimal information set at disposal here, of relative effectiveness in a dynamic setup. A bit paradoxically, audit schemes are working well if they are "deceptive": when taxpayers update their personal belief in a geographical way, implicitly assuming that the whole population is audited at random and each individual is equally likely to be picked, then the TA should optimally proceed with targeted audits, precisely because agents do not realize why they are chosen. Conversely, temporal update implies that taxpayers acknowledge that each individual is audited in way that reflects her individual features. In this situation, the introduction of some randomness on the part of the TA would shake the certainty of those who did not expected an audit and increase on average the efficacy of enforcement.

References

[Allingham and Sandmo, 1972] Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, 1(3-4):323–338.

[Alm et al., 1993] Alm, J., Cronshaw, M. B., and McKee, M. (1993). Tax compliance with endogenous audit selection rules. *Kyklos*, 46(1):27–45.

[Alm et al., 1992] Alm, J., McClelland, G. H., and Schulze, W. D. (1992). Why do people pay taxes? *Journal of Public Economics*, 48(1):21–38.

[Beyer and Schwefel, 2002] Beyer, H. and Schwefel, H. (2002). Evolution strategies - a comprehensive introduction. *Natural Computing*, 1(1):3–52.

[Bloomquist, 2003] Bloomquist, K. M. (2003). Tax evasion, income inequality and opportunity costs of compliance. Paper presented at the 96th Annual Conference of the National Tax Association.

[Collins and Plumlee, 1991] Collins, J. H. and Plumlee, R. D. (1991). The taxpayer's labor and reporting decision: The effect of audit schemes. *The Accounting Review*, 66(3):pp. 559–576.

[Kahneman and Tversky, 1979] Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):pp. 263–292.

[Kirchler et al., 2008] Kirchler, E., Hoelzl, E., and Wahl, I. (2008). Enforced versus voluntary tax compliance: The "slippery slope" framework. *Journal of Economic Psychology*, 29(2):210 – 225.

[Kleven et al., 2011] Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., and Saez, E. (2011). Unwilling or unable to cheat? evidence from a tax audit experiment in denmark. *Econometrica*, 79(3):651–692.

[Torgler, 2002] Torgler, B. (2002). Speaking to theorists and searching for facts: Tax morale and tax compliance in experiments. *Journal of Economic Surveys*, 16(5):657–683.

[Trautmann et al., 2011] Trautmann, H., Mersmann, O., and Arnu, D. (2011). http://cran.r-project.org/web/packages/cmaes/cmaes.pdf.