INTERNATIONAL MONETARY FUND

Coming Clean on Your Taxes

Sebastian Beer and Ruud de Mooij

WP/23/6

2023
JAN

WORKING PAPER

IMF Working Paper
Fiscal Affairs Department

Coming Clean on Your Taxes
Prepared by Sebastian Beer and Ruud de Mooij

Authorized for distribution by Ruud de Mooij
January 2023

ABSTRACT: This paper develops a simple model to explore whether a higher detection probability for offshore tax evaders—e.g. because of improved exchange of information between countries and/or due to digitalization of tax administrations—renders it optimal for governments to introduce a voluntary disclosure program (VDP) and, if so, under what terms. We find that if the VDP is unanticipated, it is likely to be optimal for a revenue-maximizing government to introduce a VDP with relatively generous terms, i.e. a low or even negative penalty. When anticipated, however, the VDP is neither incentive compatible nor optimal, as it induces otherwise compliant taxpayers to evade tax. A VDP can then only be beneficial if tax evasion induces an external social cost beyond the direct revenue foregone, e.g., due to adverse effects on overall tax morale. In contrast to the common view that VDPs should come along with additional enforcement effort, we find that governments should relax enforcement if the VDP itself provides more powerful incentives to come clean.

Coming Clean on Your Taxes

Prepared by Sebastian Beer and Ruud de Mooij[1]

Contents

I. Introduction

In managing the post-COVID recovery, many governments need to improve their revenue performance. Raising more taxes from the most affluent in society attracts particular interest, as the wealthy have generally fared much better through the pandemic than others, which has amplified pre-existing inequality. In this context, there is renewed interest in voluntary disclosure programs (VDPs) and tax amnesty programs (TAPs). For instance, Kenya, Nigeria, and Sri Lanka had tax amnesty programs in 2021 while Albania, Honduras, Indonesia and Trinidad and Tobago offered them in 2022. More countries might follow in the coming years.

A VDP allows previous tax evaders to disclose information about their (offshore) assets and income at reduced punitive action by government. Programs generally impose higher payment obligations than what is imposed on compliant taxpayers (through interest and penalties). However, payments and other punitive measures, such as criminal prosecution, are generally less than those for evaders who are detected. VDPs vary in terms of eligibility and payment liability. They share similarities with tax amnesty programs (TAPs), which usually forgive the full tax liability in exchange for some fixed payment. In the case of so-called 'extensive amnesties', this payment could be even lower than the principal tax (Franzoni 1996). The main goal of VDPs and TAPs is to generate additional public revenue, both in the short and medium term. In the short run, revenue gains arise from the immediate expansion of the tax base due to increased voluntary compliance. In the medium term, the voluntary disclosures could lead to more sustainable improvements in tax compliance as the tax authority can exploit the acquired information for effective enforcement in later years.[2]

While fighting tax evasion is generally a core objective of tax administrations (for an overview, see e.g., Slemrod 2007), the urgency to do so has come to increased prominence during the past decade.[3] For instance, various headline-grabbing scandals, such as the Panama papers and the Paradise papers, have elevated the fight against tax evasion on the political agenda of countries. Zucman (2013) estimates that $6 trillion of global wealth is hidden in offshore locations, and the lion share of these assets likely remains untaxed. A more recent estimate by Alstadsæter et al. (2018) finds an amount equivalent to 10 percent of world GDP. Alstadsæter et al. (2019) exploit leaked information from the Panama papers for Scandinavian individuals to show that offshore tax evasion is mostly concentrated among people at the top of the income distribution. Indeed, while around 10 percent of all households evade taxes, this share rises to between 30 and 40 percent for the top percentile (Leenders et al. 2023 report similar evidence for the Netherlands). The analysis in this paper is best understood as focusing on wealthy individuals holding offshore accounts.[4]

---

[2] Some countries have excluded the use of acquired information in subsequent tax audits (so-called closed/sealed fiscal years), which reduces the likelihood of sustained revenue increase.

[3] Not all non-compliance is willful and intentional tax evasion. Unintentional non-compliance calls for a different policy response, e.g., through taxpayer education.

[4] VDPs can also apply to corporate taxpayers, but the focus here is on individuals.

The increased awareness of tax evasion by the rich has led to important policy responses globally (Johannesen and Zucman 2014). With the establishment of the Global Forum on Transparency and Exchange of Information for Tax Purposes, countries have started to implement new global transparency and exchange of information standards to fight offshore tax evasion more effectively. Today, the Forum has 163 members, many of which have agreed to exchange a predefined set of account information held by non-residents automatically through the common reporting standards.[5] While still being implemented in many countries, automatic exchange of information (AEOI) has become operational since 2018 and has already impacted asset portfolios. For instance, a recent study by Beer et al. (2019) finds that it has reduced foreign-owned deposits in offshore jurisdictions by 25 percent. Global efforts to enhance tax transparency come at a time of rapid developments in digitalization, which are revolutionizing data management in tax administrations and help address tax evasion (see e.g., Gupta et al. 2017). In effect, these developments are significantly raising the probability of detection of offshore tax evaders, which is one of two key policy parameters (together with the penalty if caught) in the classic Allingham-Sandmo (1972) theory of tax evasion. Such effects might be most relevant for advanced economies, as developing countries generally have more limited capacity to receive and actively use information from abroad (see e.g., IMF 2022).

These trends have triggered increased interest in policies to fight tax evasion by especially wealthy individuals, such as through high-wealth individuals compliance programs (Buchanan and McLaughlin 2017). VDPs and TAPs are often considered as part of such a strategy and there is ample experience with these programs. OECD (2015) discusses 47 of them around the world. In most countries, fraudsters who are detected are obliged to pay the full amount of tax on their current and past obligations, plus interest and penalties. The penalties range widely, between 20 and 200 percent of the tax obligation. In most countries, evaders also face the possibility of criminal prosecution. Under a VDP, these punitive measures are generally significantly relaxed. For example, interest is often reduced. In 19 out of the 47 cases, penalties are waived altogether, whereas in other cases they are generally reduced significantly. In 26 cases, the criminal prosecution of evaders is waived under the VDP.[6]

Several studies have assessed the impact of VDPs and TAPs. Baer and Le Borgne (2008) discuss evidence from TAPs in several US States as well as several countries around the world (including Argentina, India, Ireland, Italy, Philippines, Turkey, and the US). Overall, they find that amnesties tend to produce mostly modest short-term gross revenue gains. For instance, amnesties imposed by US states on average had raised 0.7 percent of the relevant tax (see also Leonard and Zeckhauser 1987). However, Baer and Le Borgne also emphasize that the net revenue effects are smaller than the gross numbers (or can even be negative in the medium-term) due to increased evasion, reduced revenue from penalties, and additional administrative costs.

---

[5] The US has instead introduced the foreign account tax compliance act (FATCA), which requires foreign financial institutions to report on the foreign assets held by US citizens.

[6] Other punitive actions might still hold tax evaders from disclosing information voluntarily, such as anti-money laundering provisions or penalties imposed by the financial intelligence unit. Note that by bringing finances into the formal economy, money laundering itself can have the unexpected consequence of reducing tax evasion.

Moreover, revenue effects might be mismeasured, as amnesties often coincide with improved enforcement efforts, which might drive the actual improvements in revenue collection (Alm, McKee and Beck 1990). Baer and Le Borgne conclude that many TAPs have disappointed and that especially repeat amnesties can be counterproductive by inducing incentives for evasion by currently compliant taxpayers. Alternative policies are therefore preferred to improve compliance, such as advancements of tax administration and perhaps modification of the permanent treatment of evaders coming clean on their taxes.

A recent paper by Benedek et al. (2022) elaborates on the design, principles, and risks of VDPs. It also discusses several country experiences, including Argentina, Indonesia, Greece, Italy, Turkey, Spain, and South Africa. They list as advantages of VDPs: (i) the early revenue from base broadening; (ii) sustained revenue increases if information from the disclosures is subsequently used for compliance efforts; (iii) reduced cost of litigation/prosecution; (iv) possibly lower enforcement costs through improved voluntary compliance. However, they also point to several risks, such as (i) reduced taxpayer morale and perception of fairness and trust by rewarding evaders; (ii) increased evasion if VDPs are repeated frequently; this might be so even if only in expectation, reflecting the classic time inconsistency problem (governments cannot credibly commit to offering a VDP only once); (iii) reduced revenue from penalties compared to detecting fraudulent evaders. The paper also notes that the effectiveness of VDPs depend on legal and administrative design, including in relation to anti-money laundering and the combating financing of terrorism.

Even though VDPs are ubiquitous in practice, theoretical guidance on which to build policy prescriptions is scarce. Amid a changing global landscape of AEOI and fast-moving digitalization, however, this leaves important policy questions unaddressed by any formalized, consistent economic analysis. For instance, what conditions determine whether a VDP/TAP is desirable and effective? How do AEOI and digitalization affect this assessment? What should be the terms of a VDP if detection probabilities increase? How should the introduction of a VDP/TAP affect the enforcement effort to fight tax evasion?

A handful of studies from the early 1990s provide some but limited guidance to these questions. For instance, one important finding, emphasized by Malik and Schwab (1991), is that VDPs do not affect compliance directly. Intuitively, all else equal, a VDP will offer no benefit to evaders compared to full compliance, a choice that wasn't optimal for evaders in the first place. To make a VDP effective in inducing evaders to come clean on their taxes, there must be a shock in either their circumstances or in policies. A few studies have explored such shocks. For example, Andreoni (1990) develops a model with shocks in future consumption. This makes a permanent VDP potentially effective for tax evaders as the possibility to come clean in the future reduces uncertainty about future penalties. Indeed, evaders who are confronted with a negative consumption shock can choose to opt into the VDP to reduce the chance of being caught cheating. In Malik and Schwab (1991), there is ex-ante uncertainty because agents initially underestimate the disutility associated with evasion. As they discover that these costs are higher than anticipated, a TAP offers an opportunity to come clean. Stella (1991) explores the introduction of a temporary TAP. The paper shows this might be effective if it comes along with an

announcement of increased future enforcement effort (which is sometimes deliberately designed in actual VDP reform packages). However, as tax administrations cannot credibly commit to doing so and will rather have an incentive to exaggerate their future efforts, they find that TAPs will typically fail to have any impact on tax compliance.

In a more recent paper, Langenmeyr (2017) studies the impact of a fully anticipated VDP in a model with uncertainty about future detection probabilities. This uncertainty renders an option value to tax evasion, since agents can choose to come clean in case the detection probability turns out to be high. However, this also means that a VDP encourages otherwise compliant taxpayers to evade since they might also benefit from the VDP. Therefore, the VDP might not be optimal for the government. However, the VDP might become attractive if the detection of evaders is associated with significantly higher administrative costs for the government compared to voluntary disclosures. She provides evidence, based on a German survey among tax inspectors, that this is indeed plausible.

This paper contributes to the literature on VDPs in four important ways.[7] First, the shock in our model that gives rise to the incentive to utilize the VDP relates to an exogenous increase in the efficiency of enforcement, which subsequently raises the probability of detection. This shock is particularly interesting at the current juncture, as AEOI and new digital technologies are being implemented by countries. The shock in enforcement efficiency might also provide a plausible case under which governments could credibly commit to offer a one-time temporary VDP, without the temptation of repetition.

Second, the paper analyzes both unanticipated and anticipated VDPs in a single framework,[8] while previous studies are confined to an analysis of either of the two. In our model, agents live for two periods. In the first period, they choose between evasion and compliance based on individual tax morale (guilt from evasion) and a given probability of detection. In the second period, a shock occurs in the detection probability and the government can consider introducing a VDP to induce some previous tax evaders to come clean. If this is unanticipated, a revenue-maximizing government will find it optimal to introduce a VDP with a low or even negative penalty. However, if agents foresee the increased detection probability and the availability of a VDP in the second stage, some otherwise compliant taxpayers will start evading tax in the first period. This renders a VDP less attractive to the government and impacts the optimal design. In fact, the optimal VDP might be such that no individuals opt in. Only when we introduce a negative externality from tax evasion (reflecting, for instance, adverse effects on overall tax morale or high cost of litigation from caught evaders), an anticipated VDP can become both optimal and incentive compatible if the VDP reduces average evasion.

---

[7] The rest of the paper refers to voluntary disclosure programs but also applies to tax amnesties.

[8] We focus on the distinction between anticipated and unanticipated VDPs. This is not the same as permanent and one-off VDPs. While the former would clearly be anticipated, one-off VDPs (or their precise terms) might be unanticipated. However, some anticipation might occur if the one-off VDP is deemed attractive for the government to impose or if VDPs are repeated frequently.

As a third contribution, the paper explores the relationship between an optimally designed VDP and optimal enforcement effort by the tax administration.[9] Previous studies have generally assumed that VDPs come along with increased enforcement effort and thus pictured these programs as strategic complements to tax enforcement. However, this is not necessarily the optimal response. By incentivizing voluntary disclosure of tax evaders, VDPs may also reduce the marginal return to enforcement activity and allow the government to save on its administrative effort. Indeed, we find that, if the government uses a lower penalty in the VDP to encourage tax evaders to come clean, optimal detection efforts will also decrease (thus capitalizing part of the exogenous efficiency improvement in detection).

A final contribution of the paper is that it presents simulations based on parameters from existing tax systems and plausible enforcement probabilities. These numerical results provide further intuition and policy guidance on the use and design of VDPs.

The rest of this paper is organized as follows. The next section develops the basic model of offshore tax evasion as a discrete choice and derives optimal enforcement. Then, we extend the basic model to two periods, with the introduction of a VDP in the second stage. First, we assume that the VDP is unanticipated by individuals and derive the optimal VDP penalty in the presence of a shock in detection probabilities, as well as the optimal enforcement effort. Subsequently, we do the same for a VDP that is anticipated and where individuals can adjust their behavior prior to its introduction. The final section concludes.

---

[9] For more about optimal tax administration, see Keen and Slemrod (2017).

II. Basic Model

We develop a simple model for a continuum of wealthy individuals, each holding the same stock of offshore assets that yields a periodic return y, which is taxable in the residence country at a tax rate τ. Individuals are risk-neutral and take a discrete decision as to whether they evade the tax or report income to the government and pay tax. This choice depends on their heterogeneous guilt level and the policy parameters set by the government, such as the tax rate, penalties in case of detected evasion and the enforcement effort that steers the probability of detection. Optimal policy is subsequently derived by maximizing government revenue, net of enforcement costs.

A. Tax Evasion

Risk neutral individuals take a discrete decision whether to evade taxes or not. If compliant, they receive net of tax capital income $(1 - \tau)y$, which is also their utility:

$$u^c = (1 - \tau)y \tag{1}$$

If they do not report their income, evasion is detected with probability p and individuals caught concealing assets pay a fine $\varphi > 1$ times the tax due. With probability $(1 - p)$, evasion is not detected. However, utility from concealed income is only $(1 - g_i)y$, where g_i is a guilt parameter that differs between individuals. Individual i's expected utility from evasion is thus equal to:[10]

$$u_i^e(p) = [1 - p\tau\varphi - (1 - p)g_i]y \tag{2}$$

---

[10] Langenmeyer (2017) discusses the case of risk-averse agents. Expressions become more complicated in that case, but key results with respect to the VDP effects carry over in her model – as is the case in ours (available on request). In contrast to Langenmeyer, guilt in our model reduces taxpayer's utility only if evasion is not detected. The model's predictions remain qualitatively unchanged if guilt reduces utility independently from the tax administration's detection success.

Figure 1. Utility from Evasion Versus Compliance

$(1 - p\varphi\tau)y$

$u_i{}^e$

$(1 - \tau)y$

u^c

g_i

0               $\bar{g}$ 1
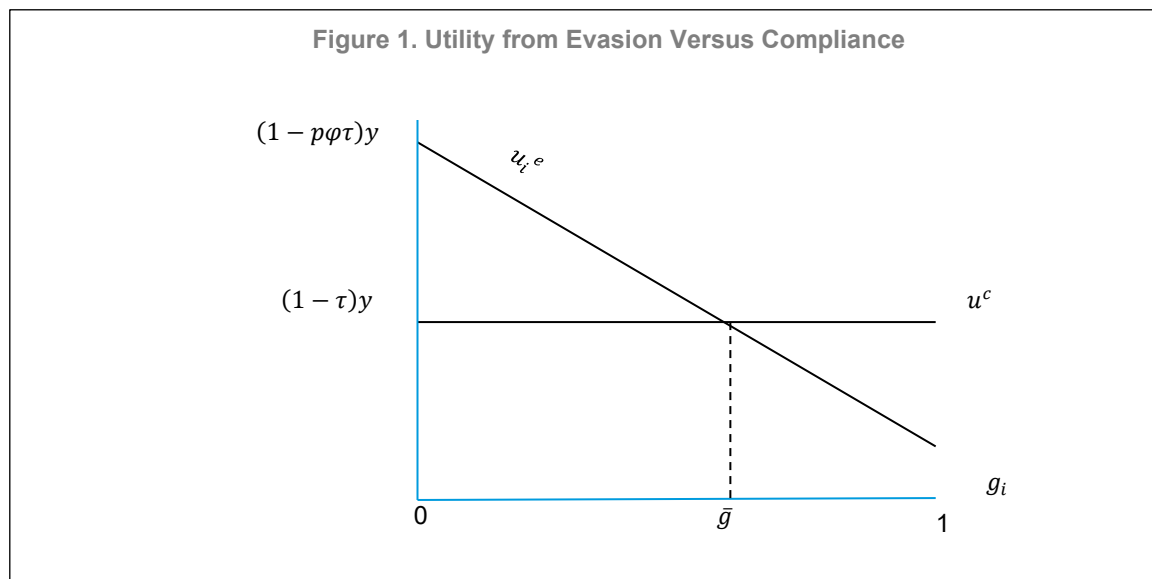
Hence, utility from evasion is a decreasing function of guilt, which we assume is distributed uniformly over the unit interval (which normalizes population size to unity). As utility from compliance is constant across individuals, there exists (at most) one level of guilt, denoted by $\bar{g}$, at which individuals are indifferent between compliance and evasion. This is illustrated in Figure 1, which shows utility associated with the two options as a function of the individual guilt parameter, g_i. In case of full tax compliance, utility is independent of guilt and hence represented by the horizontal line in the graph. Expected utility in case of evasion is a declining function of the guilt parameter. The intersection of the two lines reflects the threshold level $\bar{g}$. Wealthy individuals with lower guilt levels prefer evasion, while those with a guilt above the threshold prefer voluntary compliance. The steady state threshold, given by

$$\bar{g}(p) = \tau \frac{1 - p\varphi}{1 - p} \tag{3}$$

is a decreasing function of the detection probability and the penalty. If the government imposes no penalty in case of detected evasion ($\varphi = 1$), only taxpayers for whom the non-pecuniary guilt exceeds the tax rate ($g_i > \tau$) will truthfully report their income. The tax rate τ thus provides a natural upper bound for $\bar{g}$. For $\varphi > 1$, the threshold will be lower. For instance, if we assume $\varphi = 3$, $\tau = 0.3$ and $p = 0.1$, we obtain $\bar{g} = 0.23$. The uniform distribution of guilt over the unit interval implies that $\bar{g}$ reflects the share of evaders in the population and $(1 - \bar{g})$ is the share of compliant taxpayers.[11] Hence, in the numerical example 23 percent of individuals choose evasion.

---

[11] Both types of taxpayers exist only if $p < \frac{1}{\varphi}$. If p exceeds this level, there will be no evaders. For example, if $\varphi = 3$ (i.e., a fine of 200 percent, which should be interpreted as including the expected cost of prosecution), p must be smaller than 0.33.

B. Optimal Tax Administration

Revenue in the steady state derives from two sources: a share $(1 - \bar{g})$ of taxpayers pays τy while a share $\bar{g}$ contributes $\varphi \tau y$ with a probablity of p. We assume the government maximizes revenue collection from offshore accounts net of administrative costs $C(p, \bar{g})$ by choosing an optimal detection probability:[12]

$$\max_{p} R = [(1 - \bar{g}) + \bar{g} p \varphi - C(p, \bar{g})] \tau y \qquad (4)$$

Administrative costs are given by

$$C(p, \bar{g}) = \frac{\lambda}{2} p^2 + \alpha \bar{g} \qquad (5)$$

where $\alpha \geq 0$ and $\lambda > 0$ are non-negative parameters. The first term of the cost function captures the simple insight that implementing a higher detection probability p requires an increase in resources spent on enforcement activity. We assume that this relationship between costs and detection is convex, so that additional resources deployed towards detection come with a diminishing marginal return. The parameter λ represents the administration's enforcement (in)efficiency in detecting tax evaders.

The second term captures a negative externality from tax evasion, measured by the parameter α. It suggests that enforcement costs increase in the share of tax evaders, $\bar{g}$. This can have alternative interpretations: One is that the cost of collecting tax from detected evaders is higher than for taxpayers who voluntarily comply, for instance due to litigation costs. Another interpretation is that, as taxpayers are conditionally cooperative (Fischbacher and others, 2001), public coffers suffer from the presence of evaders indirectly through its impact on overall tax morale (including through evasion of other taxes). As will become clear later, it is this negative externality from evasion that can render VDPs optimal.

The first-order condition[13] for an optimal detection probability reads

$$\bar{g}(p)\varphi - \frac{\partial \bar{g}(p)}{\partial p}[1 - p\varphi] = \frac{dC(p, \bar{g})}{\partial p} \qquad (6)$$

Expression (6) equates the marginal revenue of a rise in p on the left-hand side to the marginal cost of enforcement to increase p on the right-hand side. The first term on the left measures the direct revenue increase from detecting evasion of non-compliant taxpayers, of which there is a share $\bar{g}$, at a rate that includes

---

[12] In contrast, we treat the penalty in case of detection as an exogenous parameter that is determined by the wider prevailing legal framework. Revenue maximization is a reasonable approximation of welfare maximization if society assigns a low weight to the utility of wealthy individuals with offshore accounts.

[13] The second-order condition, given by $-2\frac{\tau(\varphi-1)}{(1-p)^2}\left(\varphi + \frac{\alpha}{1-p}\right) - \lambda < 0$, is satisfied for all parameter values.

the full penalty on evaders. The second term on the left captures a deterrence effect: some evaders will start to comply in response to a marginal increase in the detection probability, an effect that is measured by $\frac{\partial \bar{g}}{\partial p} = \frac{-\tau(\varphi - 1)}{(1-p)^2} < 0$. Accounting for the expected penalty previous evaders will cease to pay, the marginal revenue raised from such taxpayers is $1 - p\varphi$.

C. Improvement in Detection Efficiency

The optimal detection probability increases in detection efficiency. To see this, implicitly differentiate equation (6) with respect to λ and use the definitions of $\bar{g}$ and C from equations (3) and (5), which implies

$$-\frac{\partial p^*}{\partial \lambda} = \frac{p^*}{\lambda + 2\frac{\tau(\varphi - 1)}{(1 - p)^2}\left(\varphi + \frac{\alpha}{1 - p}\right)} > 0 \tag{7}$$

Accordingly, there is a positive association between the efficiency of enforcement (measured by a lower λ) and the optimal probability of detection (p^*). Yet, the optimal detection probability rises less than proportional in λ: defining the elasticity of the optimal detection probability as $\varepsilon \equiv -\frac{\partial p^*}{\partial \lambda}\frac{\lambda}{p^*}$, it follows that $0 < \varepsilon < 1$.

To explore the implications of a higher detection efficiency on the administrative effort by the government, we totally differentiate $C(p, \bar{g})$ to get

$$\frac{dC(p^*, \bar{g})}{d\lambda} = \left(\frac{1}{2} - \varepsilon\right)p^2 + \alpha\frac{\tau(\varphi - 1)}{(1 - p)^2}\frac{p\varepsilon}{\lambda} \tag{8}$$

The first term on the right-hand side of (8) shows that an improvement in enforcement efficiency ($d\lambda < 0$) has two opposing effects on the optimal budget of the revenue administration. On the one hand, it directly reduces the cost of existing enforcement efforts so that the administrative budget can be reduced. This effect is equal to $-\frac{1}{2}p^2$. On the other hand, a lower λ induces the government to increase the detection probability, which increases enforcement costs (an effect that is measured by εp^2). On balance, the administrative budget will increase because of improved enforcement efficiency if $\varepsilon > \frac{1}{2}$; otherwise, the tax administration budget will shrink. The second term on the right-hand side of (8) modifies this effect through the negative impact of a higher detection probability on the share of tax evaders, which reduces enforcement costs through α.

D. Simulations

For a range of parameter values, Table 1 shows numerical values of the optimal detection probability (p^*), the equilibrium share of tax evaders ($\bar{g}$), the elasticity ε, government net revenue in percent of potential revenue

$(\frac{R}{\tau y})$, and total enforcement cost in percent of revenue $\frac{C\tau y}{R}$. The simulations assume a tax rate $\tau = 0.3$, negative externalities of evasion costing $\alpha = \{0.2, 0.4\}$, a penalty $\varphi = \{2\frac{1}{2}, 4\}$ and a detection efficiency $\lambda = \{15, 30\}$.

Table 1 shows that the optimal detection probability rises in the efficiency of detection (decreases in λ), but less than proportional as we know from (7) and from the column showing that ε is smaller than one. For instance, for $\varphi = 2\frac{1}{2}$ and $\alpha = 0.2$, the probability of detection rises less than twofold from 4.1 percent to 7.9 percent if λ drops from 30 to 15. The optimal probability also increases in the penalty for evaders and in the indirect enforcement cost of evasion. The associated share of tax evaders ranges between 17.8 percent if the fine, enforcement efficiency and indirect cost of evasion are high, and 28.1 percent when they are all low. Revenue net of enforcement cost is 76.6 percent of potential gross revenue if the fine is high and direct and indirect enforcement costs are low; revenue decreases to 61.1 percent if the fine is low and enforcement costs are high. Enforcement costs range between 12.3 and 24.9 percent of tax revenue.

Table 1. Optimal Detection Probability and Share of Tax Evaders in Steady State

Parameters				Outcomes			
λ	φ	α	p^*	ε	$\bar{g}$	$\frac{R}{\tau y}$	$\frac{C\tau y}{R}$
15	2.5	0.2	0.08	0.91	0.26	0.69	0.14
30	2.5	0.2	0.04	0.96	0.28	0.67	0.12
15	4	0.2	0.11	0.68	0.19	0.77	0.17
30	4	0.2	0.06	0.83	0.24	0.71	0.15
15	2.5	0.4	0.09	0.92	0.26	0.64	0.25
30	2.5	0.4	0.04	0.96	0.28	0.61	0.23
15	4	0.4	0.12	0.69	0.18	0.73	0.24
30	4	0.4	0.07	0.84	0.23	0.67	0.25

Note: Solution for p^* using equation (6) with $\tau = 0.3$.

III. Unanticipated Voluntary Disclosure Program

We now expand the basic model to a two-period setting to analyze the introduction of a VDP in period 2. We first consider an unanticipated VDP, whereby taxpayers do not foresee the availability of a VDP in the second period when deciding to evade taxes in the first period. Behavior in the first period is therefore given by the steady state equilibrium derived in the previous section. In the second period, the government introduces a VDP that allows first-period evaders to come clean at a reduced penalty $v < \varphi$. In a real-world setting, this unanticipated VDP case can guide the policy choice of a one-off unannounced VDP. For analytical convenience, we assume that audit outcomes in the first period do not affect eligibility of VDP participation. Hence, if taxpayers evaded tax in the first period, they may still participate in the VDP in the second period, irrespective of whether they were caught evading or not.[14] Appendix I relaxes this assumption by restricting VDPs only to evaders who were not caught in the first period.

A. Tax Evasion and VDP Participation

Figure 2 illustrates the utility space in period 2 as a function of the guilt parameter, g_i, assuming a constant detection probability over time ($p_2 = p_1 = p$). As in Figure 1, the horizontal line with intercept $(1 - \tau)$ is utility from compliance. The downward sloping line is utility from evasion. Individuals with guilt below $\bar{g}$ evade taxes in period 2, while those with g_i between $\bar{g}$ and 1 comply (as they do in period 1).

Figure 2. Utility in period 2 if $p_2 = p_1 = p$
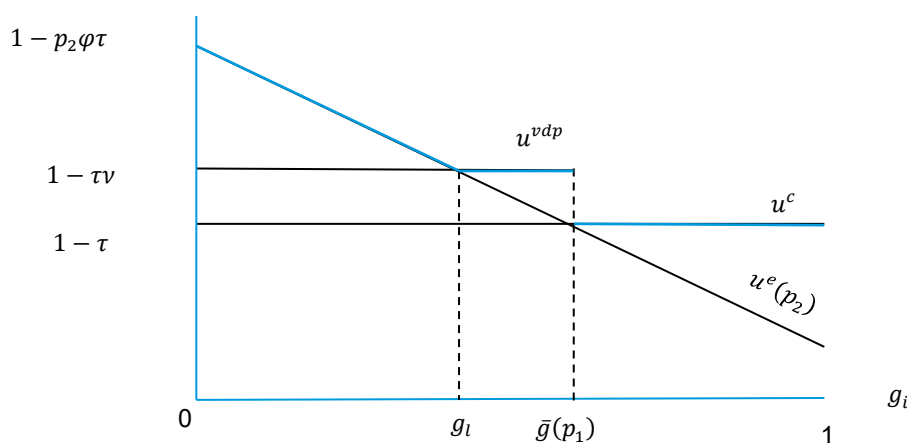
Now, suppose the government introduces a VDP in period 2. Utility from entering the VDP is represented by the horizontal line with intercept $(1 - \tau v)$, reflecting the VDP penalty v. Previous evaders would be willing to

---

[14] All decisions are taken ex-ante based on expectations and actual outcomes ex post do not affect behavior in our model.

enter the VDP if utility in the program exceeded that of continued evasion. As taxpayers with guilt below $\bar{g}$ preferred evasion over compliance in the first period, utility in the VDP needs to exceed that of compliance i.e., $(1 - v\tau) > (1 - \tau)$, which can only happen if $v < 1$. Hence, the VDP effectively induces taxpayers to come clean only if it provides a discount relative to the standard tax regime for compliant taxpayers – reminiscent to the extensive amnesty referred to in the introduction.

In what follows, we say a VDP is incentive compatible if some individuals find it attractive to enter the program. To characterize this condition, we define g_l as the guilt level for which utility from entering the VDP, measured by $u^{vdp} = (1 - v\tau)y$, equals the expected utility from continued evasion in the second period, measured by $u_i^e(p_2) = [1 - p_2\tau\varphi - (1 - p_2)g_i]y$. Solving for the guilt parameter gives:

$$g_l = \frac{\tau(v - p_2\varphi)}{1 - p_2} \tag{9}$$

Where the subscript l indicates that the guilt parameter is a lower bound. Individuals with a guilt parameter below this threshold still choose to evade.

Definition: An unanticipated VDP is incentive compatible only if $g_l < \bar{g}(p_1)$, where $\bar{g}(p_1)$ is given by equation (3) with $p = p_1$ and g_l is defined in equation (9).

Incentive compatibility requires that the mass of first-period evaders entering the VDP, measured by $\bar{g}(p_1) - g_l(p_2)$, is positive. It is a function of the detection probabilities in both periods and the effective penalty in the VDP. The condition for incentive compatibility can also be expressed in terms of the VDP penalty by using equations (3) and (9):

$$v < 1 + \frac{(p_2 - p_1)(\varphi - 1)}{1 - p_1} \equiv v^{res\_u} \tag{10}$$

Where we denote v^{res_u} the reservation penalty for the unanticipated VDP, defined as the maximum VDP penalty that still renders the program incentive incompatible.

Equation (10) shows that when the detection probability remains constant between periods ($p_2 = p_1$), a VDP would only attract participants if the program penalty is smaller than one – the reservation penalty in this case (as also illustrated in Figure 2). Conversely, when the detection probability increases, incentive compatible VDPs will exist with $v^{res\_u} > 1$. Figure 3 illustrates the utility space for such a scenario. Relative to a constant detection probability, utility from evasion is now lower and more sensitive to changes in taxpayer-specific guilt, reflected by the steeper declining slope. Even though utility in the VDP is lower than utility from compliance, the program will attract previous evaders with guilt levels between g_l and $\bar{g}$, as the higher detection probability renders continued evasion a less attractive option for them.

Figure 3. Utility in Period 2 if $p_2 > p_1$
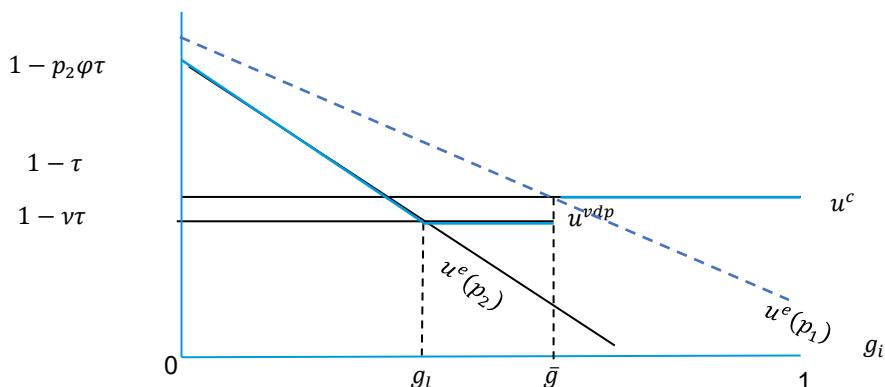
B. Optimal Policy Under Constant Detection Efficiency

The government maximizes revenue with respect to enforcement effort in period 2 and the VDP penalty. The VDP penalty can be interpreted as an instrument to set a distinct (and optimal) effective tax rate, $v\tau$, for a subset of taxpayers in the second period (while the general tax rate is fixed). In what follows, we first assume that the efficiency of detection remains unchanged between the first and second period. Subsequently, we examine the optimal policy response to an increasing efficiency of detection, represented by a drop in λ in period 2.

Second-period revenue derives from three sources: (i) a share $1 - \bar{g}(p_1)$ of individuals is compliant in period 2 and pays taxes honestly; (ii) a fraction $\bar{g}(p_1) - g_l$ of period-1 evaders enters the VDP in period 2 and pay an effective tax rate of τv; (iii) a fraction g_l of individuals continue to evade tax in period 2 and pay an expected penalty of $\varphi \tau p_2$. Revenue net of administrative costs in the second period thus reads as:

$$R_2 = [1 - \bar{g}(p_1) + [\bar{g}(p_1) - g_l]v + \varphi p_2 g_l - C(p_2, g_l)]\tau y \qquad (11)$$

The first-order condition with respect to p_2 determines the optimal detection probability and is given by:

$$g_l \varphi - \frac{\partial g_l}{\partial p_2}[v - p_2 \varphi] = \frac{dC(p_2, g_l)}{\partial p_2} \qquad (12)$$

It resembles the first-order condition of the basic model in (6) (and thus optimal enforcement in period 1) with two differences: (i) the net revenue effect from reducing evasion is now $v - p_2 \varphi$ (rather than $1 - p_1 \varphi$) due to the

VDP penalty; and (ii) the share of evaders is g_l (rather than $\bar{g}$). Equation (12) also implies that, if the VDP penalty equals one, the optimal detection probability remains unchanged between the first and second period.[15]

The first order condition for the optimal VDP penalty reads as:

$$[\bar{g}(p_1) - g_l] - \frac{\partial g_l}{\partial v}(v - \varphi p_2) = \alpha \frac{\partial g_l}{\partial v} \tag{13}$$

It requires that the additional revenue from a higher VDP penalty, on the left-hand side, is equated to additional enforcement costs, on the right. The additional revenue depends first on the share of VDP adopters who will pay the VDP fine, captured by $\bar{g}(p_1) - g_l$; and second, on the decline in the number of VDP adopters if the VDP penalty is raised, which causes a net revenue loss of $v - \varphi p_2$. Indirect enforcement costs, measured by α, are also associated with a marginal reduction in the number of VDP adopters, increasing the pool of second-period evaders.

Using the definition of the lower bound guilt threshold (equation (9)) and $\frac{\partial g_l}{\partial v} = \frac{\tau}{1-p_2}$, the first-order condition in (13) can be rewritten as

$$g_l = \frac{1}{2}\left[\bar{g}(p_1) - \frac{\alpha \tau}{1 - p_2}\right] \tag{14}$$

Hence, in the absence of negative externalities from evasion ($\alpha = 0$), the optimal VDP is characterized by a very simple rule, namely, that the penalty should be set such that the share of tax evaders in period 2 will be exactly half the share of tax evaders in period 1. This rule is independent from the probability of detection in either of the two periods and the implied VDPs are incentive compatible as $g_l < \bar{g}(p_1)$ holds in all cases. When evasion imposes additional external costs ($\alpha > 0$), equation (14) shows that the optimal share of evaders is only further reduced.

Using the definitions for $\bar{g}$ and g_l in (3) and (9) and rearranging terms, we can rewrite (14) also as an explicit expression for the optimal VDP fine

$$v^* = \frac{1}{2}\left[1 + \frac{(p_2 - p_1)(\varphi - 1)}{1 - p_1} + p_2 \varphi - \alpha\right] = \frac{1}{2}[v^{res\_u} + p_2 \varphi - \alpha] \tag{15}$$

---

[15] This follows from noting that $g_l = \bar{g}(p_2)$ when $v = 1$.

Hence, v^* is half the reservation penalty with two further adjustments: (i) there is an upward effect that depends on the expected penalty paid by continued evaders in period 2 ($p_2\varphi$); and (ii) there is a negative effect that depends on the external cost of evasion (α).[16]

To explore the relationship between the optimal detection probability and the optimal VDP penalty, we rewrite the first-order condition (12) by substituting $\frac{\partial g_l}{\partial p_2} = -\frac{(\varphi-v)g_l}{(1-p_2)(v-p_2\varphi)}$ and rearrange terms as:

$$\Phi^{period\,2} \equiv g_l\left[\frac{\varphi-v}{1-p_2} + \varphi\right] + \alpha\tau\frac{\varphi-v}{(1-p_2)^2} - \lambda_2 p_2 = 0 \qquad (16)$$

Implicitly differentiating equation (16) with respect to v yields an expression for the change in the optimal detection probability in response to the VDP penalty:[17]

$$\frac{\partial p_2}{\partial v} = -\frac{\Phi_v^{period\,2}}{\Phi_{p_2}^{period\,2}} > 0, \qquad (17)$$

Hence, in addressing tax evasion, the government faces a trade-off between the carrot and the stick, i.e., the carrot of a generous VDP penalty that may induce many previous evaders to come clean at relatively attractive terms; and the stick of tougher enforcement efforts, which increase the probability of detection in case of continued evasion. Indeed, equation (17) suggests that a high VDP penalty will make it attractive for the government to rely on a higher detection probability to induce evaders to come clean and pay the higher fines. Conversely, a low VDP penalty will by itself attract a larger share of previous tax evaders so that enforcement efforts can be scaled back, and government can save on administrative costs.

If the detection efficiency remains unchanged, it is optimal for the government to set a negative penalty in the VDP scheme, i.e., previous tax evaders who come clean face a lower effective tax burden than compliant taxpayers. In response to the low VDP penalty, the government will find it optimal to also reduce the detection probability and thus save on administrative costs. To see this, substitute equation (14) into equation (16) to reflect the optimal joint policy mix in period 2 and compare with equation (6) for the optimal detection probability in period 1. Assuming $\alpha = 0$, these two equations imply the following: $p_2 < p_1 \iff \frac{1}{2\lambda_2}\left[\frac{\varphi-v^*}{1-p_2} + \varphi\right] <$

---

[16] The revenue maximizing VDP penalty in (15) is most likely smaller than one, unless the detection probability rises sufficiently (as explored below). Hence, the government may find it attractive to use an extensive amnesty where it imposes a lower tax burden on previous evaders than on compliant taxpayers. Intuitively, as government restricts the VDP to previous tax evaders without extending it to previously compliant taxpayers, its revenue maximizing strategy is to offer a reduced effective tax rate (through a negative penalty rate in the VDP) to attract more evaders to come clean.

[17] Where the inequality follows from noting that $\Phi_{p_2}^{period\,2} = \frac{\partial\Phi^{period\,2}}{\partial p_2} = \frac{-2\tau(\varphi-v)}{(1-p_2)^3}(\varphi-v-\alpha) < 0$ if equation (12) characterizes a

maximum and $\Phi_v^{period\,2} = \frac{\partial\Phi^{period\,2}}{\partial v} = \frac{2\tau\left(\varphi-v-\frac{\alpha}{2}\right)}{(1-p_2)^2} > 0$ follows as a result.

$\frac{1}{\lambda_1}\left[\frac{\varphi-1}{1-p_1}+\varphi\right]$. Together with the optimal penalty in (15), this implies that $p_2 < p_1$ if the detection efficiency is unchanged ($\lambda_2 = \lambda_1$). [18]

Note, however, that a lower effective tax rate for previous evaders than for compliant taxpayers might not be possible for moral and political reasons. If the government restricts the VDP scheme to those where $v \geq 1$, our results imply that it will be optimal to set $v = 1$. In that case, the optimal detection probability in period 2 will remain unchanged relative to the first period and equation (10) implies that the VDP scheme is just at its reservation point. The restricted optimal policy ($v \geq 1$) therefore resembles the steady state in period 1 where the VDP is simply an allowance for previous evaders to come clean at the same tax terms as all-time compliers.

C. Optimal Policy Under Increasing Detection Efficiency

We next explore the optimal policy response when the efficiency of detection in the second period increases relative to the first period. Implicitly differentiating equations (13) and (16) with respect to λ gives the marginal response in the two policy parameters

$$\frac{\partial v^*}{\partial \lambda_2} = \frac{1}{2}\left[\frac{\varphi-1}{1-p_1}+\varphi\right]\frac{\partial p_2}{\partial \lambda_2} \tag{18.a}$$

$$\Phi_{p_2}^{period\ 2}\frac{\partial p_2}{\partial \lambda_2} + \Phi_{v}^{period\ 2}\frac{\partial v^*}{\partial \lambda_2} - p_2 = 0 \tag{18.b}$$

Equation (18.a) shows that the change in the optimal VDP penalty in response to a higher detection efficiency has the same sign as the change in the second-period enforcement effort. Indeed, if the probability of detection increases, so will the optimal penalty under the VDP scheme, as a way to raise revenue.

If the VDP penalty were fixed, equation (18.b) implies that the response in the optimal detection probability to a marginal increase in detection efficiency is given by $\frac{-p_2}{\Phi_{p_2}^{period\ 2}} > 0$, resembling the response in the steady state. If the VDP penalty can be modified as part of the re-optimization (and if it is not constrained by any lower bound, such as $v \geq 1$), we see that the impact on the detection probability is even exacerbated (see Appendix II):

$$-\frac{\partial p_2}{\partial \lambda_2} = \frac{-p_2}{\Phi_{p_2}^{period\ 2} + \frac{1}{2}\left[\frac{\varphi-1}{1-p_1}+\varphi\right]\Phi_{v}^{period\ 2}} > \frac{-p_2}{\Phi_{p_2}^{period\ 2}} > 0 \tag{19}$$

---

[18] To see this, note from (16) that $\frac{\varphi-v^*}{1-p_2}+\varphi = 2\varphi - \frac{1}{2}\left[\frac{1-p_1\varphi}{1-p_1}-\frac{\alpha}{1-p_2}\right]$. Setting $\alpha = 0$ and substituting in equation (17), we obtain $\varphi - \frac{1}{4}\frac{(1-p_1\varphi)}{1-p_1} < \varphi + \frac{\varphi-1}{1-p_1}$, which is obviously satisfied.

In other words, the availability of a VDP increases the sensitivity of enforcement efforts to changes in detection efficiency.

Even though the detection probability and the optimal penalty increase in response to efficiency shocks, the optimal VDP fine may remain below unity. For an optimal unanticipated VDP to have a positive penalty – or, at least, not to reduce the effective tax rate below that of all-time compliers – there needs to be a sufficiently large increase in the detection efficiency. From the definition of the optimal penalty (equation 15), we see that $v^* \geq 1$ requires that

$$p_2 - p_1 \geq \frac{1 - p_1 \varphi + \alpha}{\frac{\varphi - 1}{1 - p_1} + \varphi}, \qquad (20)$$

which is larger if α is large and φ and p_1 are small. For example, the required increase in the detection probability is 9 percentage points if $\varphi = 4$, $p_1 = 0.2$ and $\alpha = 0.5$, i.e., the probability of detection would need to increase from 20 to 29 percent. If $\varphi = 2.5$, $p_1 = 0.1$ and $\alpha = 1$, it would have to increase by 42 percentage points, i.e., from 10 to 52 percent.

If condition (20) does not hold and we restrict $v^* \geq 1$, the government cannot implement its optimal revenue maximizing VDP. However, the government might still find it attractive to introduce a VDP at the constraint $v = 1$. Appendix III shows that this indeed unambiguously raises revenue.

D. Simulations

Table 2 presents numerical simulations for the optimal policy in period 2 under the same parameter configurations as in Table 1. For λ_2, we assume a value of 3. This implies a reduction by 90 percent if $\lambda_1 = 30$ and by 80 percent if $\lambda_1 = 15$.

The improved detection efficiency in period 2 unambiguously increases the detection probability, reduces the number of tax evaders, and raises revenue; it reduces administrative costs in most cases. The optimal effective penalty under the VDP scheme in period 2 ranges from 0.65 to 1.39, i.e., the optimal effective tax paid by previous evaders can be either higher or lower than the effective tax for all-time compliers.[19]

The increase in the detection probability depends on whether the detection efficiency increases moderately ($\lambda_1 = 15$) or extensively ($\lambda_1 = 30$). If detection is more costly in period 1, the optimal detection probability is initially lower and the share of evaders higher. In that case, the government finds it optimal in period 2 to boost enforcement effort by a lot to raise the detection probability. This allows for a relatively higher VDP penalty. In

---

[19] Appendix III shows optimal detection efforts if the penalty is restricted to be non-negative.

comparison, if detection efficiency increases more moderately, the detection probability rises less in period 2. Instead, the government relies on a lower VDP penalty to induce evaders to come clean. In other words, if improvements in detection efficiency are modest, the government will combine a more generous carrot of a low VDP penalty with a more modest stick of a low detection effort to address evasion in period 2, relative to the case of a larger increase in detection efficiency.

Table 2. Optimal Unanticipated VPD and Enforcement Under Alternative Parameters

Parameters									Outcomes			
						Unconstrained υ^*						
λ_1	λ_2	φ	α	p_1^*	p_2^*	υ^*	$\bar{g}(p_1)$	$g_l(p_2)$	$\dfrac{R_1}{\tau y}$	$\dfrac{R_2}{\tau y}$	$\dfrac{C_1 \tau y}{R_1}$	$\dfrac{C_2 \tau y}{R_2}$
15	3	2.5	0.2	0.08	0.20	0.75	0.26	0.09	0.69	0.83	0.14	0.09
30	3	2.5	0.2	0.04	0.21	0.80	0.28	0.10	0.67	0.83	0.12	0.11
15	3	4	0.2	0.11	0.25	1.12	0.19	0.06	0.77	0.91	0.17	0.11
30	3	4	0.2	0.06	0.30	1.39	0.24	0.08	0.71	0.93	0.15	0.17
15	3	2.5	0.4	0.09	0.20	0.65	0.26	0.05	0.64	0.82	0.25	0.10
30	3	2.5	0.4	0.04	0.22	0.71	0.28	0.06	0.61	0.81	0.23	0.12
15	3	4	0.4	0.12	0.24	0.97	0.18	0.01	0.73	0.91	0.24	0.10
30	3	4	0.4	0.07	0.30	1.28	0.23	0.03	0.67	0.91	0.25	0.16

Note: Solution for p_2^* using equation (16) and for υ^* expression (15), with $\tau = 0.3$.

IV. Anticipated Voluntary Disclosure Program

A concern with the unanticipated VDP is that it might not be realistic to assume that rational individuals do not foresee its introduction. Indeed, sometimes a VDP is a permanent feature of a country's policy framework so that it will clearly be foreseen to stay. Moreover, if a government has a history of using unannounced temporary VDPs, individuals may expect such a policy will be implemented in the future. But even if there is no such history, individuals may nevertheless anticipate that the government will introduce a VDP if it is known that such a policy will maximize government revenue. This section therefore considers the case where individuals anticipate the government's policies in period 2, i.e., both the VDP introduction and the increased detection efficiency. Individuals can adjust their behavior in the first period to account for the possibility of later entering the VDP. In turn, the government can take these anticipation effects into account when deciding about its policies.[20]

A. Tax Evasion and VDP Participation

Suppose individuals foresee in period 1 both the probability of detection in period 2 and the introduction of a VDP. Taxpayers will then have three options during the two periods: (i) they comply with the tax in both periods and enjoy (twice) net of tax income (U^c); (ii) they evade the tax in both periods, with expected utility (U_i^e) depending on the detection probabilities in the two periods and the individual guilt parameter; (iii) they evade the tax in the first period, where evasion is detected with probability p_1, and enter the VDP in the second period (U^{vdp}), where the detection probability is p_2. Ignoring discounting between periods, utility from these three options is given by, respectively,[21]

$$U^c = 2u^c \tag{21.a}$$

$$U_i^e(p_1, p_2) = 2u_i^e(\bar{p}) \tag{21.b}$$

$$U^{vdp}(p_1) = u_i^e(p_1) + (1 - \tau v)y \tag{21.c}$$

where $\bar{p} = \frac{1}{2}(p_1 + p_2)$ is the average detection probability across the two periods. Note that if the government decides not to introduce a VDP, equation (21.c) is not feasible and the optimal choice between evasion and compliance is determined by (21.a) and (21.b) alone. This takes us back to the model of the steady state above, but with $p = \bar{p}$. Hence, individuals foresee the future change in the detection probability and determine

---

[20] Appendix V extends our two-period model to a T-period setup, whereby the VDP in period 2 is followed by T-2 periods in which those who opt in for the VDP are treated as fully compliant taxpayers in subsequent years. The conditions for incentive compatibility and optimality change slightly and the specific choice of period 2 parameters becomes less important as the horizon expands (and in the absence of discounting).

[21] Utility from evading in both periods is $U^e = (1 - p_1\varphi\tau - (1 - p_1)g_i) + (1 - p_2\varphi\tau - (1 - p_2)g_i)$, which can be rewritten as $2 - \varphi\tau(p_1 + p_2) - g_i(2 - p_1 - p_2) = 2u_i^e(\bar{p})$. This assumes that evaders who get caught in period 1 can choose to evade again in period 2. If this option is ruled out, e.g., as the detection in period 1 makes it harder to conceal income in period 2, expressions are slightly modified, as shown in Appendix I. The same applies to evaders who are caught in period 1 and who choose to opt in the VDP in period 2 – for which Appendix I also looks at the case of excluding this as an option.

their behavior in both periods on the basis of the average probability. This is because there is no opportunity for them to switch and come clean on their taxes in period 2. This differs from the model without anticipation effects, where the choice in period 1 is determined only by p_1.

Figure 4 illustrates the utility associated with the three options. Building on Figure 1, the horizontal line represents utility under full tax compliance in both periods, while expected utility in case of evasion in both periods is the steeply declining line. In case of evasion in period 1 and entering the VDP in period 2, utility is also represented by a decreasing function of guilt. However, the slope is less steep than under full evasion since entering the VDP in the second period relieves taxpayers from the moral cost associated with guilt. In Figure 4, the curve of the VDP option intersects with both the utility under full evasion and utility under full compliance, reflecting guilt levels at which individuals are indifferent between the VDP option and either of these alternatives.

Figure 4. Payoffs from Entering an Anticipated VDP
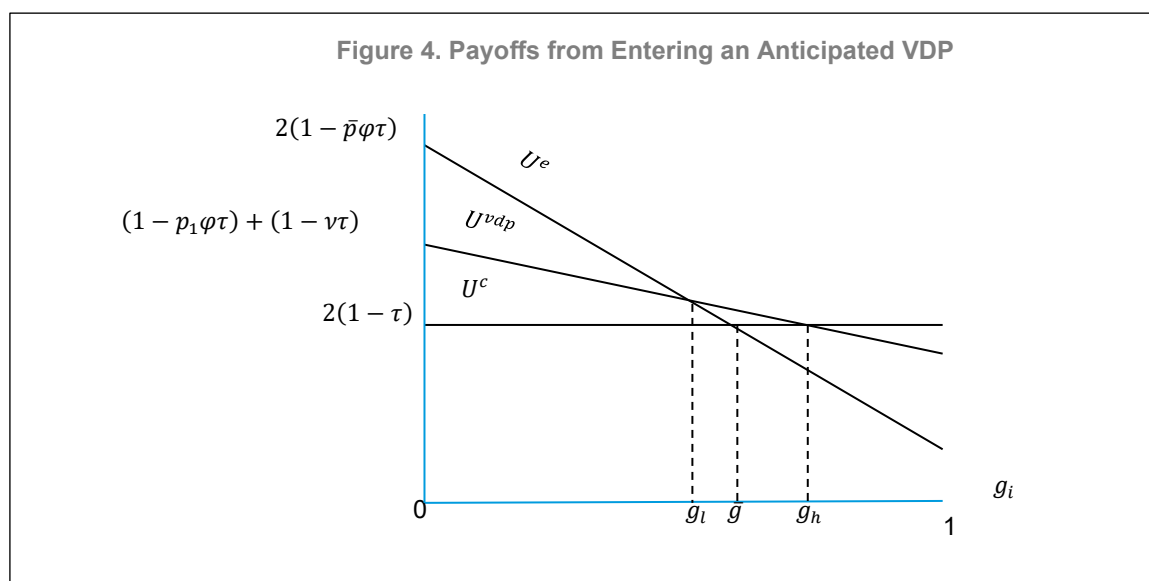
More formally, we first define g_l as the guilt level for which utility from entering the VDP equals the utility from evasion in both periods: $U^{vdp}(p_1) = U_i^e(p_1, p_2)$. Solving for the guilt parameter gives:

$$g_l = \frac{\tau(v - p_2\varphi)}{1 - p_2} \tag{22}$$

which is the same as (9). As shown in Figure 4, individuals with a guilt parameter below this threshold will evade in both periods.[22]

---

[22] This is the case illustrated in Figure 4 but does not necessarily hold for all parameters. For instance, if the penalty under the VDP is sufficiently high, the two curves intersect beyond $\bar{g}$, in which case some taxpayers below g_l would prefer compliance over evasion.

Second, we define g_h as the guilt level for which utility from entering the VDP equals the utility from full tax compliance in both periods: $U^{vdp}(p_1) = U^c$. Solving for the guilt parameter suggests that this second threshold is given by:

$$g_h = \frac{\tau(\rho - p_1\varphi)}{1 - p_1} \tag{23}$$

where $\rho = 2 - v$. As shown in Figure 4, taxpayers with guilt levels exceeding g_h will enjoy the highest utility when they fully comply with the tax in both periods. For guilt levels between g_l and g_h, they will enjoy the highest utility if they evade in period 1 and enter the VDP in period 2.

Individuals who opt for the VDP thus consist of two groups. The first group are individuals with guilt levels between g_l and $\bar{g}(\bar{p})$: they would have chosen to evade tax in the absence of a VDP, but the VDP induces them to come clean in the second period. This is generally the group that VDPs aim to target. The second group consists of individuals with guilt levels between $\bar{g}(\bar{p})$ and g_h: they would have chosen to comply in period 1 if there were no VDP but are induced to evade tax in anticipation of the VDP in the second period. The VDP thus generates increased tax evasion in period 1, which is an unintended effect of the VDP. The dividing line among the VDP opt-ins between previous period 2 evaders and previous period 1 compliers is guilt level $\bar{g}(\bar{p})$, at which individuals would be indifferent between evasion and compliance in the absence of a VDP.

As before, the anticipated VDP is said to be incentive compatible if at least some individuals will opt into the program.

Definition: An anticipated VDP is incentive compatible if only if $g_l < g_h$, where the threshold guilt levels are defined in (22) and (23) respectively.

Using (22) and (23), we show that this inequality is satisfied when:

$$v < 1 + \frac{1}{2}\frac{(p_2 - p_1)}{1 - \bar{p}}(\varphi - 1) = v^{res\_a} \tag{24}$$

Where we call v^{res_a} the reservation penalty in case of the anticipated VDP, defined as the maximum level of the VDP penalty that is incentive compatible.[23] Condition (24) shows that, if the detection probability remains constant over time ($p_2 = p_1$), the reservation penalty equals one ($v^{res\_a} = 1$), as with the unanticipated VDP. Hence, there will be no VDP with a positive penalty that can induce individuals to come clean.[24] If the detection

---

[23] Note also that $v^{res\_a} \leq v^{res\_u}$, i.e., the condition for a VDP with $v > 1$ to be incentive compatible is less stringent if it is unanticipated compared to the case of anticipation.

[24] Of course, a VDP with $v < 1$ can be incentive compatible (as in the previous section).

probability increases between period 1 and period 2, condition (24) shows that a VDP with a positive penalty can be incentive compatible ($v^{res\_a} > 1$). Hence, rising probabilities are a necessary condition for a VDP with a positive penalty to be effective in encouraging evaders to come clean. The feasible penalty rate under the VDP cannot exceed the reservation penalty in (24). For example, if $\varphi = 3$, $p_2 = 0.2$ and $p_1 = 0.1$, we obtain $v^{res\_a} = 1.11$, i.e., with a doubling of the probability of detection, the maximum penalty under which a VDP can be incentive compatible is only 11 percent.

B. Optimal VDP

Next, we explore whether a VDP will be optimal for the government in maximizing its revenue. Government revenue collected over the two periods, net of collection costs, is equal to:

$$R = [2(1 - g_h) + (g_h - g_l)v + (p_1 g_h + p_2 g_l)\varphi - C(p_1, g_h) - C(p_2, g_l)]\tau y \qquad (25)$$

The first term on the right-hand side of (25) reflects the fraction $(1 - g_h)$ of individuals who voluntarily comply in both periods and pay twice the full tax τy. The second term in (25) measures the fraction $(g_h - g_l)$ in period 2 of previous evaders who enter the VDP and pay the effective tax $\tau v y$. The third term in (25) captures the penalties from detected tax evaders in both periods, paying effectively $\varphi \tau y$: in period 1, this applies to the remaining share g_h of evaders who are detected with probability of p_1; in period 2, a fraction g_l continues to evade and is detected with probability p_2.

The government maximizes net revenue by choosing the penalty in the VDP (v) and the detection probabilities in both periods (p_1 and p_2). The first order condition for the optimal penalty in the VDP is given by:

$$(g_h - g_l) - \frac{\partial g_h}{\partial v}(\rho - \varphi p_1) - \frac{\partial g_l}{\partial v}(v - \varphi p_2) = \alpha[\frac{\partial g_h}{\partial v} + \frac{\partial g_l}{\partial v}] \qquad (26)$$

A marginal increase in the VDP penalty has two effects on government revenue. First, it directly increases revenue from participants of the VDP, of which there is a mass of $g_h - g_l$. Second, a higher VDP penalty reduces the incentive for individuals to enter the scheme, which applies to the two groups of VDP entrants identified above. The first group are the otherwise compliant taxpayers who decide to evade tax in the first period to be able to enter the VDP. The higher penalty discourages such a strategy for a fraction of $-\frac{\partial g_h}{\partial v} = \frac{g_h}{\rho - p_1 \varphi} > 0$ of the population. For them, a higher VDP raises more revenue at the margin, which is given by $(\rho - \varphi p_1)$. The second group are those who otherwise evade tax in both periods. A higher penalty under the VDP discourages this share of the population by $\frac{\partial g_l}{\partial v} = \frac{g_l}{v - \varphi p_2} > 0$ and for them revenue raised changes by $-(v - \varphi p_2)$.

Substituting the partial derivatives, and using the expressions for g_l and g_h in (22) and (23), the first-order condition can be rewritten as:

$$2(g_h - g_l) = \alpha\tau\left(\frac{p_2 - p_1}{(1 - p_1)(1 - p_2)}\right) \tag{27}$$

Expression (27) shows that, if there are no external costs of tax evasion on tax morale or litigation (i.e., if $\alpha = 0$), we obtain an optimally chosen VDP penalty such that it satisfies $g_h = g_l$. This means that the optimal VDP is not incentive compatible, i.e., it is designed such that no individual chooses to join the VDP. In terms of Figure 4, utility under the VDP scheme, U^{vdp}, intersects the other two lines exactly at $\bar{g}$. Hence, there is no reason for the government to introduce a VDP.

Only if the share of tax evaders raises the cost of tax collection ($\alpha > 0$), equation (27) suggests that there can be an optimal VDP that satisfies $g_h > g_l$, i.e., there exist a VDP that is incentive compatible. Moreover, such a VDP can be optimal only if the detection probability increases in the second period, i.e., $p_2 > p_1$.

Substituting the expressions for g_l and g_h into (27) and solving for the optimal fine gives

$$v^{**} = 1 + \frac{1}{2}\frac{(p_2 - p_1)}{1 - \bar{p}}\left(\varphi - 1 - \frac{\alpha}{2}\right) = v^{res\_a} - \frac{1}{4}\frac{p_2 - p_1}{1 - \bar{p}}\alpha \tag{28}$$

Hence, $\alpha = 0$ implies that the optimal fine under the VDP is equal to its reservation level, $v^{**} = v^{res\_a}$, implying that no individual will participate in the VDP. Also, recall from equation (24) that $v^{res\_a} = 1$ if the detection probability remains unchanged across the two periods ($p_2 = p_1 = \bar{p}$). If both $\alpha > 0$ and $p_2 > p_1$, equation (28) implies that $v^{**} < v^{res\_a}$, i.e., the government will set the optimal VDP fine below the reservation penalty to induce some evaders to come clean in order to reduce the external cost of evasion. Whether the optimal penalty under the VDP is positive will depend on the sign of the term $\varphi - 1 - \frac{\alpha}{2}$. For example, if $\varphi = 3$, $\alpha = 2$, $p_2 = 0.2$ and $p_1 = 0.1$, we obtain $v^{**} = 1.06$. For $\alpha > 4$, however, we would obtain $v^{**} < 1$.

C. Optimal Administration

We next derive the optimal detection probabilities in the two periods in the presence of a VDP. Government revenue net of enforcement costs is given by (25) and the first order conditions for maximizing this with respect to p_1 and p_2 can be expressed as:[25]

$$\Phi^1(p_1, v, \lambda_1) \equiv g_h\left[\frac{\varphi - \rho}{1 - p_1} + \varphi\right] + \alpha\tau\frac{\varphi - \rho}{(1 - p_1)^2} - \lambda_1 p_1 = 0 \tag{29.a}$$

---

[25] Here, we use the definitions in (22) and (23) and substitute the derivatives $\frac{\partial g_h}{\partial p_1} = -\tau\frac{\varphi - \rho}{(1 - p_1)^2} < 0$ and $\frac{\partial g_l}{\partial p_2} = -\tau\frac{\varphi - v}{(1 - p_2)^2} < 0$.

$$\Phi^2(p_2, v, \lambda_2) \equiv g_l \left[\frac{\varphi - v}{1 - p_2} + \varphi \right] + \alpha\tau \frac{\varphi - v}{(1 - p_2)^2} - \lambda_2 p_2 = 0 \qquad (29.b)$$

When the efficiency of detection does not change between periods, an optimal solution is characterized by $v^{**} = 1$ and $p_1^* = p_2^*$, implying that the revenue maximizing strategy is to set the penalty at the reservation rate and leave the detection probability unchanged. This is easily verified by recognizing that, when $v = 1$, $g_h = g_l$ and $\rho = v$ so that (29) can only hold for $p_1 = p_2$. In turn, when the detection probability remains unchanged, equation (28) suggests that the optimal fine is $v^{**} = 1$. Hence, $v^{**} = 1$ and $p_1^* = p_2^*$ simultaneously satisfy (28) and (29). This result is markedly different from that under an anticipated VDP. Indeed, while the optimal unanticipated VDP offers a discounted effective tax rate to increase VDP participation of previous evaders and reduce the cost of enforcement, the optimal anticipated VDP does not offer such discount since it would trigger more evasion in period 1. The result is that none of the previous tax evaders will opt in to the VDP.

To obtain a better understanding of the optimal detection probabilities when the detection efficiency increases, we use a linear approximation of the decision rules in (28) and (29) around the identified solution $\bar{p} = p_1 = p_2$ and at $v = 1$. This yields (see Appendix IV):

$$p_2^* - p_1^* \approx \frac{\varepsilon}{\lambda} \left[\bar{p}(\lambda_1 - \lambda_2) - 2\Phi_v^1(v - 1) \right] \qquad (30)$$

where $\varepsilon \equiv -\frac{\partial \bar{p}}{\partial \lambda} \frac{\lambda}{p} > 0$ defined before lies between 0 and 1 and $\Phi_v^1 = -\frac{2\tau}{(1-\bar{p})^2} \left(\varphi - 1 - \frac{\alpha}{2} \right) < 0$. Equation (30) yields two important insights. First, an increase in the detection efficiency, represented by $\lambda_1 - \lambda_2 > 0$, will increases the optimal detection probability. The size of this effect is measured by $\frac{\varepsilon\bar{p}}{\lambda}$. Second, equation (30) shows that the VDP penalty has an unambiguous positive impact on the optimal detection probability. Moreover, the optimal VDP penalty increases in the difference between first and second period detection probabilities (equation 28), suggesting that the optimal solution satisfies $p_2^* > p_1^*$ and $v^* > 1$. As with the unanticipated VDP, the administration's ability to adjust the optimal VDP penalty in the face of increasing enforcement efficiency also raises the sensitivity of the optimal detection response.

An interesting question is whether a VDP reduces the average number of tax evaders across the two periods. In the absence of a VDP, the share of tax evaders is the same over the two periods and given by $2\bar{g}(\bar{p})$.[26] With a VDP, the share of evaders in period 1 is $g_h(p_1)$ which is higher than without a VDP; in period 2, it equals $g_l(p_2)$, which is lower than without a VDP. On balance, we find that the average number of evaders across the two periods is lower with the VDP, as long as the VDP penalty is strictly below the reservation penalty.[27]

D. Simulations

Table 3 shows numerical simulations for the optimal VDP penalty and the optimal detection probabilities when (28) and (29) hold simultaneously. The simulations assume the same variation in parameters as in Tables 1 and 2 and shows the optimal policies both if there is no VDP and if there is a VDP. As in Table 2, the parameter for detection efficiency declines from 30 or 15 to 3.

Improvements in enforcement efficiency in period 2 increase the optimal detection probability. For instance, in the first row of Table 3, the first-period detection probability is 6 percent and this increases to 32 percent in period 2. In the bottom row, the detection probability increases ninefold, from 4 to 37 percent. In the absence of a VDP, the share of tax evaders is the same across the two periods.

If there is a VDP, Table 3 shows that the optimal fine under the VDP is positive for all parameter configurations. In the first row, for instance, υ^{**} is set at 1.2, i.e., the fine is 20 percent of the tax liability. The highest optimal penalty is 50 percent in the bottom row. Similarly, detection probabilities increase in the second period along with the improvement in detection efficiency. In the first row, for instance, the detection probability increases from 7 percent in the first period to 32 percent in the second.

---

[26] In the absence of a VDP, the first order conditions for an optimal detection rate are

$$\Phi^{1,no-vdp}(p_1, v, \lambda_1) \equiv \bar{g}\left[\frac{\varphi - 1}{1 - \bar{p}} + \varphi\right] + \alpha\tau\frac{\varphi - 1}{(1 - \bar{p})^2} - \lambda_1 p_1 = 0$$

$$\Phi^{2,no-vdp}(p_2, v, \lambda_2) \equiv \bar{g}\left[\frac{\varphi - 1}{1 - \bar{p}} + \varphi\right] + \alpha\tau\frac{\varphi - 1}{(1 - \bar{p})^2} - \lambda_2 p_2 = 0$$

[27] To contrast the number of evaders in the presence of a VDP with the number of evaders without a VDP, note that $g_l(p_2) = g_h(p_1) = \bar{g}(\bar{p})$ when the VDP fine is set at $v^{res} = \frac{1-p_2}{1-\bar{p}} + \frac{p_2-p_1}{1-\bar{p}}\frac{\varphi}{2}$. Furthermore, both guilt thresholds are linear functions of the VDP fine. We thus get $g_l(p_2) = \bar{g}(\bar{p}) + (v - v^{res})\tau\left(\frac{1}{1-p_2}\right)$ and similarly $g_h(p_1) = \bar{g}(\bar{p}) - (v - v^{res})\tau\left(\frac{1}{1-p_1}\right)$. Adding the equalities, we get $g_h + g_l = 2\bar{g} + (v - v^{res})\tau\left(\frac{1}{1-p_2} - \frac{1}{1-p_1}\right)$. Since $v^* < v^{res}$, we obtain that $g_h + g_l < 2\bar{g}$ when detection probabilities increase.

Table 3. Optimal Anticipated VPD and Enforcement Under Alternative Parameters

Parameters							Outcomes					
λ_1	λ_2	φ	α	p_1^*	p_2^*	υ^{**}	g_h	g_l	$\dfrac{R_1}{\tau y}$	$\dfrac{R_2}{\tau y}$	$\dfrac{C_1 \tau y}{R_1}$	$\dfrac{C_2 \tau y}{R_2}$
Without VDP												
15	3	2.5	0.2	0.06	0.32		0.19	0.19	0.77	0.77	0.09	0.26
30	3	2.5	0.2	0.03	0.33		0.20	0.20	0.76	0.76	0.07	0.27
15	3	4	0.2	0.06	0.32		0.09	0.09	0.89	0.85	0.05	0.20
30	3	4	0.2	0.03	0.34		0.10	0.10	0.88	0.84	0.04	0.22
15	3	2.5	0.4	0.07	0.36		0.18	0.18	0.75	0.72	0.15	0.37
30	3	2.5	0.4	0.04	0.37		0.19	0.19	0.74	0.71	0.13	0.39
15	3	4	0.4	0.07	0.35		0.06	0.06	0.90	0.82	0.07	0.25
30	3	4	0.4	0.04	0.37		0.07	0.07	0.89	0.80	0.05	0.29
With VDP												
15	3	2.5	0.2	0.07	0.32	1.20	0.20	0.18	0.76	0.98	0.10	0.19
30	3	2.5	0.2	0.03	0.33	1.23	0.21	0.19	0.75	1.01	0.08	0.20
15	3	4	0.2	0.07	0.31	1.42	0.10	0.08	0.87	1.25	0.06	0.13
30	3	4	0.2	0.04	0.33	1.50	0.11	0.08	0.86	1.33	0.05	0.13
15	3	2.5	0.4	0.07	0.35	1.19	0.20	0.15	0.71	0.98	0.17	0.24
30	3	2.5	0.4	0.04	0.35	1.22	0.22	0.15	0.70	1.01	0.15	0.25
15	3	4	0.4	0.08	0.33	1.42	0.09	0.04	0.86	1.28	0.09	0.14
30	3	4	0.4	0.04	0.35	1.50	0.10	0.04	0.85	1.37	0.08	0.15

Note: Solution using equations (28) and (29)), with $\tau = 0.3$.

Compared to the case without a VDP, the optimal detection probability in the first period is slightly higher in the presence of a VDP. The optimal detection probability in the second period, in contrast, tends to be slightly lower. Intuitively, the VDP is used to attract evaders to come clean, which allows the government to save on administrative costs by reducing enforcement efforts. As a result, in the presence of a VDP, the share of tax evaders in period 1 is higher while it is lower in period 2 – reflecting the VDP entrants. The share of VDP entrants varies in the simulations: in the first row, it is only 2 percent while in the last row the share of evaders declines from 10 to 4 percent.

Interestingly, government revenue in period 2 exceeds the revenue under full compliance in some of the scenarios. The reason is that the penalty payments under the VDP regime boost revenue beyond the level under full compliance.

V. Conclusion

This paper explores whether improvements in the detection of offshore tax evaders (e.g., due to improved exchange of information between countries and/or digitalization of tax administration) can make it attractive for governments to adopt a VDP and at what terms. We find that, if individuals do not anticipate future policies, a one-off VDP can be attractive for governments to maximize their tax revenue from offshore wealth. To induce previous tax evaders to come clean on their taxes, the VDP will need to offer a low or even negative penalty.

However, if a one-off VDP is such an attractive policy, wealthy individuals might foresee its introduction. In that case, we find that a VDP will be neither optimal nor effective in reducing tax evasion due to anticipation effects. Indeed, otherwise compliant taxpayers are induced to evade tax prior to its introduction to subsequently become eligible for the VDP. The VDP thus reflects a classic example of a time inconsistent policy. Only if tax evasion imposes an external cost on society that goes beyond the direct revenue foregone (such as costs from reduced overall tax morale or cost of litigation), we find that a VDP that attracts previous evaders can become effective and efficient.

While our model is highly stylized and does not capture some of the real-world complexities in designing VDPs and other tax compliance policies, the analysis in this paper offers conceptual guidance regarding the conditions and trade-offs that policy makers face. Overall, the analysis shows that the conditions for a VDP to be socially beneficial are rather stringent and depend on the rise in detection probabilities, anticipation effects, and external costs from evasion. The model helps explain why many VDPs in the past have been ineffective in improving tax compliance, e.g. as penalties in the VDP might have been set too high or detection probabilities might not have sufficiently increased. Moreover, the analysis also illustrates the trade-off governments face in incentivizing offshore tax evaders to come clean, namely between the carrot of low penalties in the VDP and the stick of higher enforcement efforts.

The paper adds to a small literature on the economic impact of VDPs, which needs further elaboration and analysis to guide policy. For instance, trends in information exchange and digitalization of tax administrations call for empirical analysis of how this interacts with the adoption of VDPs. Moreover, more empirical analysis is needed to better understand the impact of VDPs (of lack thereof) on tax evasion, revenue and enforcement, as well as how these relations are shaped by the theory. Finally, several assumptions in our model could be generalized, such as risk neutrality of agents, the separable cost function, and the uniform distribution of guilt. Special design features of VDPs could also be explored, e.g., how they relate to other enforcement efforts by e.g. financial intelligence units and anti-money laundering efforts.

Appendix I. Optimal VDP penalty if detected evaders are forced to comply in period 2

This appendix derives the optimal VDP penalty in case tax evaders who get detected in period 1 are forced to comply with the tax in period 2. This contrast with the model in the main text where all period-1 evaders can choose to either evade tax or enter the VDP in period 2. The assumption here might be more realistic as hiding income from the tax authority might have become impossible after detection. The results for the optimal penalty under the VDP are either the same (for the unanticipated) or very similar (for the anticipated), although analytically somewhat more complicated. Analytical results for the optimal detection probability are not derived here, as they become analytically too cumbersome. Numerical simulations (available upon request) indicate that results are very similar though, as compared to the model in the main text.

A. Unanticipated VDP

As in the main body of the paper, suppose the share $\bar{g}$ denotes individuals who decide to evade in the first period. Evaders who were not detected, still face the following options:

$$U^{vdp} = (1 - \tau v)$$

$$U^e = p_2(1 - \tau\varphi) + (1 - p_2)(1 - g_i)$$

Equality of the utilities implies that taxpayers with guilt levels above

$$g_l = \frac{\tau(v - p_2\varphi)}{1 - p_2}$$

will participate in the VDP. Since this is the same as in the main text, the reservation penalty in Eq. (10) remains unchanged.

A fraction $p_1\bar{g}$ of them is detected evading. In contrast to the main text, suppose that these detected evaders need to pay the full tax of a compliant taxpayer in the second period. Revenue (net of administrative costs) in the second period is

$$R_2 = 1 - \bar{g} + (\bar{g} - g_l)v + g_l\varphi p_2 + p_1\{\bar{g} - (\bar{g} - g_l)v - g_l\varphi p_2\} - C(p_2, g_l)$$

where the last term between curly brackets reflects the revenue from evaders who have been detected in period 1 and who will comply with their tax obligation in period 2.

First period evasion is $\bar{g}$ and second period evasion is $g_l(1 - p_1)$. Using $\frac{dg_l}{dv} = \frac{\tau}{1-p_2}$, the marginal effect on enforcement costs in period 2 is thus

$$\frac{\partial C}{\partial v} = \frac{\alpha\tau}{1 - p_2}$$

The first-order condition for an optimal penalty in the VDP is

$$\frac{\partial R_2}{\partial v} = (\bar{g} - g_l)(1 - p_1) - \frac{\partial g_l}{\partial v}(v - p_2\varphi)(1 - p_1) = \alpha\frac{\partial g_l}{\partial v}(1 - p_1)$$

which is identical to Eq. (13) and thus implies the same optimal penalty as in Eq. (15). Hence, while revenue expands under this alternative assumption, optimal policy rules are unchanged.

B. Anticipated VDP

Under the conditional approach where period-1 evaders who got detected are forced to comply in period 2, utility for the three strategies (C,C), (E,E) and (E,V) are as follows:

$$U^c = 2(1 - \tau)$$

$$U^e = (1 - p_1)\big[(1 - p_2)(1 - g_i)2 + p_2[(1 - g_i) + (1 - \varphi\tau)]\big] + p_1[(1 - \tau\varphi) + (1 - \tau)]$$
$$= 2 - p_1\tau - \varphi\tau(p_1 + p_2 - p_1p_2) - g_i(2(1 - p_1) - p_2 + p_1p_2)$$

$$U^{vdp} = (1 - p_1)[(1 - g_i) + (1 - \tau v)] + p_1[(1 - \tau\varphi) + (1 - \tau)]$$

The corresponding threshold for indifference between U^c and U^e is

$$\bar{g} = \frac{\tau(2 - p_1 - \varphi(p_1 + p_2 - p_1p_2))}{(1 - p_1)(2 - p_2)}$$

For indifference between U^e and U^{vdp} it is

$$g_l = \frac{\tau(v - p_2\varphi)}{1 - p_2}$$

And for indifference between U^c and U^{vdp} it is

$$g_h = \frac{\tau\big(2 - v(1 - p_1) - p_1(\varphi + 1)\big)}{1 - p_1}$$

As in the main text, we can derive the reservation penalty as the level of v that is just incentive compatible, i.e. which ensures that $g_h > g_l$. This level is given by:

$$v^{res} = 1 + \frac{p_2 - p_1}{(1 - p_1)(2 - p_2)}(\varphi - 1)$$

This penalty rate is unambiguously larger than in Eq. (24) in the main text as long as $p_2 > p_1$ since the denominator of the ratio on the right-hand side of the equation is smaller as long as $p_1(1 - p_2) > 0$, which is always true. For instance, if $p_1 = 0.1$, $p_2 = 0.3$, $\varphi = 2.5$ and $\alpha = 0.5$, the reservation penalty is 1.196 compared to 1.188 according to Eq. (24). Hence, the VDP is more likely to be incentive compatible.

Revenue also changes. In the first and second period, revenue is

$$R_1 = (1 - g_h) + p_1 g_h \varphi - C(p_1, g_h)$$

$$R_2 = (1 - g_h) + (g_h - g_l)v + g_l \varphi p_2 + p_1\{g_h - (g_h - g_l)v - g_l \varphi p_2\} - C(p_2, g_l)$$

Where the term between curly brackets summarizes adjustments made as some who evaded where caught and pay all taxes instead of the expected penalty or the VDP fine.

The first-order condition for the fine in the VDP on total revenue is

$$\frac{\partial R_1 + R_2}{\partial v} = (1 - p_1)(g_h - g_l) - \frac{\partial g_h}{\partial v}(2 - p_1\varphi - v(1 - p_1) - p_1) - \frac{\partial g_l}{\partial v}(v - \varphi p_2 - vp_1 + \varphi p_1 p_2)$$

$$= \alpha\left(\frac{\partial g_h}{\partial v} + \frac{\partial g_l}{\partial v}(1 - p_1)\right)$$

Note that $\frac{dg_h}{dv} = -\tau$ and $\frac{dg_l}{dv} = \frac{\tau}{1 - p_2}$. Using this, we can rewrite the above as

$$2(g_h - g_l) = \alpha\tau\frac{p_2 - p_1}{(1 - p_1)(1 - p_2)}$$

which is exactly the same rule as Eq. (27) in the main text. Using the above definitions and rearranging, we get for the optimal penalty

$$v^{**} = v^{res} - \alpha\frac{p_2 - p_1}{2(1 - p_1)(2 - p_2)} = 1 + \frac{p_2 - p_1}{(1 - p_1)(2 - p_2)}\left(\varphi - 1 - \frac{1}{2}\alpha\right)$$

The optimal penalty is larger than the one in Eq. (28) of the main text as long as $p_2 > p_1$. In the numerical example above, for instance, the optimal penalty is 1.163, which is larger than 1.156 according to Eq. (28). Intuitively, a low VDP penalty becomes less effective to induce evaders to come clean if a portion of them is no longer eligible to participate.

Appendix II. Impact of higher detection efficiency on detection probability with unanticipated VDP

The denominator of equation (21) is given by

$$\Phi_{p_2}^{period\ 2} + \frac{1}{2}\left(\frac{\varphi - 1}{1 - p_1} + \varphi\right)\Phi_{\nu}^{period\ 2}$$

Substituting the partial derivatives $\Phi_{p_2}^{period\ 2} = -2\frac{\tau(\varphi - \nu)}{(1 - p_2)^2}\left(\varphi + \frac{\alpha}{1 - p_2}\right) - \lambda_2$ and $\Phi_{\nu}^{period\ 2} = \frac{2\tau}{(1 - p_2)^2}\left(\varphi - \nu - \frac{\alpha}{2}\right)$, and rearranging terms, we get:

$$-\lambda_2 - \frac{2\tau}{(1 - p_2)^2}\left[(\varphi - \nu)\frac{1 - p_1\varphi}{2(1 - p_1)} + \alpha\left[\frac{\varphi - \nu}{1 - p_2} + \frac{1}{4}\left(\frac{\varphi - 1}{1 - p_1} + \varphi\right)\right]\right] < 0$$

which is negative and decreasing in α.

Appendix III. Revenue effect of an unanticipated constrained VDP

To verify whether a VDP with $v = 1$ increases government revenue, we compare net revenue in period 2 under the no-VDP case (with $\bar{g}(p_1)$ evaders) with net revenue under a VDP with penalty $v = 1$:

$$R_2^{no} = \left[1 - \bar{g}(p_1) + \varphi p_2 \bar{g}(p_1) - C(p_2, \bar{g}(p_1))\right]\tau y$$

$$R_2^{v=1} = \left[1 - g_l + \varphi p_2 g_l - C(p_2, g_l)\right]\tau y$$

Subtracting the former from latter yields the additional revenue from introducing a VDP with $v = 1$:

$$R_2^{v=1} - R_2^{no} = \left[\bar{g}(p_1) - g_l\right](1 - \varphi p_2 + \alpha)\tau y > 0$$

Hence, an unanticipated VDP with $v = 1$ will be revenue-increasing for the government if it is incentive compatible, i.e., if $\bar{g}(p_1) > g_l$ – which is always the case if $p_2 > p_1$. And when the detection efficiency increases, the optimal detection probability increases as well at $v = 1$. A higher detection efficiency is thus a necessary condition for implementing a revenue maximizing VDP that imposes a non-negative penalty on previous evaders.

The table below illustrates the optimal outcomes numerically if we restrict $v \geq 1$, using the same parameters as in Table 2 of the main text.

									Constrained $v = 1$			
λ_1	λ_2	φ	α	p_1^*	p_2^*	v^*	$\bar{g}(p_1)$	$g_l(p_2)$	$\dfrac{R_1}{\tau y}$	$\dfrac{R_2}{\tau y}$	$\dfrac{C_1 \tau y}{R_1}$	$\dfrac{C_2 \tau y}{R_2}$
15	3	2.5	0.2	0.08	0.26	1	0.26	0.14	0.691	0.821	0.143	0.162
30	3	2.5	0.2	0.04	0.26	1	0.28	0.14	0.667	0.821	0.123	0.162
15	3	4	0.2	0.11	0.22	1	0.19	0.05	0.766	0.913	0.165	0.090
30	3	4	0.2	0.06	0.22	1	0.24	0.05	0.713	0.913	0.153	0.090
15	3	2.5	0.4	0.09	0.29	1	0.26	0.11	0.639	0.795	0.247	0.220
30	3	2.5	0.4	0.04	0.29	1	0.28	0.11	0.611	0.795	0.231	0.220
15	3	4	0.4	0.12	0.24	1	0.18	0.01	0.729	0.907	0.244	0.102
30	3	4	0.4	0.07	0.24	1	0.23	0.01	0.666	0.907	0.249	0.102

Note: Solution for p_1^*, p_2^* and v^* using equation (15) and (20)), with $\tau = 0.3$. $v^*$is should be at least one.

Appendix IV. Linear approximation of the optimal detection probability with anticipated VDP

The two first order conditions for optimal detection probabilities are given by equations (29.a) and (29.b) in the main text. We can approximate these functions with a first-order Taylor expansion around the point $(\bar{p}, 1, \lambda)$ where $\bar{p}$ is an average detection rate and λ the average detection efficiency:

$$\Phi^1(p_1, \nu, \lambda_1) \approx \Phi^1(\bar{p}, 1, \lambda) + \Phi_p^1(\bar{p}, 1, \lambda)(p_1 - \bar{p}) + \Phi_\nu^1(\bar{p}, 1, \lambda)(\nu - 1) - \bar{p}(\lambda_1 - \lambda)$$

$$\Phi^2(p_2, \nu, \lambda_2) \approx \Phi^2(\bar{p}, 1, \lambda) + \Phi_p^2(\bar{p}, 1, \lambda)(p_2 - \bar{p}) + \Phi_\nu^2(\bar{p}, 1, \lambda)(\nu - 1) - \bar{p}(\lambda_2 - \lambda)$$

The partial derivatives are given by

$$\Phi_p^1(\bar{p}, 1, \lambda) \equiv \frac{\partial \Phi^1(\bar{p}, 1, \lambda)}{\partial p_1} = -2\frac{\tau(\varphi - 1)}{(1 - \bar{p})^2}\left(\varphi + \frac{\alpha}{1 - \bar{p}}\right) - \lambda < 0$$

$$\Phi_\nu^1(\bar{p}, 1) \equiv \frac{\partial \Phi^1(\bar{p}, 1, \lambda)}{\partial \nu} = -\frac{2\tau}{(1 - \bar{p})^2}\left(\varphi - 1 - \frac{\alpha}{2}\right)$$

$$\Phi_p^2(\bar{p}, 1, \lambda) \equiv \frac{\partial \Phi^2(\bar{p}, 1, \lambda)}{\partial p_2} = -2\frac{\tau(\varphi - 1)}{(1 - \bar{p})^2}\left(\varphi + \frac{\alpha}{1 - \bar{p}}\right) - \lambda < 0$$

$$\Phi_\nu^2(\bar{p}, 1) \equiv \frac{\partial \Phi^2(\bar{p}, 1, \lambda)}{\partial \nu} = \frac{2\tau}{(1 - \bar{p})^2}\left(\varphi - 1 - \frac{\alpha}{2}\right)$$

For the second order condition to hold for all parameter values, we need $\Phi_p^1(\bar{p}, 1, \lambda) < 0$ and $\Phi_p^2(\bar{p}, 1, \lambda) < 0$, which is the case. The sign of $\Phi_\nu^1(\bar{p}, 1, \lambda)$ and $\Phi_\nu^2(\bar{p}, 1, \lambda)$ depends on the value of α relative to φ, as is the case in equation (16). If $\varphi - 1 - \frac{\alpha}{2} > 0$ (which is the condition found in the main text for ν to exceed one), we have $\Phi_\nu^1(\bar{p}, 1, \lambda) < 0$ and $\Phi_\nu^2(\bar{p}, 1, \lambda) > 0$

Note that $\Phi^1(\bar{p}, 1, \lambda) = \Phi^2(\bar{p}, 1, \lambda)$, $\Phi_\nu^1(\bar{p}, 1, \lambda) = -\Phi_\nu^2(\bar{p}, 1, \lambda)$, and $\Phi_p^1(\bar{p}, 1, \lambda) = \Phi_p^2(\bar{p}, 1, \lambda)$. We thus get

$$\Phi^1(p_1, \nu, \lambda) - \Phi^2(p_2, \nu, \lambda) \approx 2\Phi_\nu^1(\bar{p}, 1)(\nu - 1) + \Phi_p^1(\bar{p}, 1, \lambda)(p_1 - p_2) + \bar{p}(\lambda_2 - \lambda_1) = 0$$

Rearranging gives

$$p_2 - p_1 = \frac{\bar{p}(\lambda_2 - \lambda_1) + 2\Phi_v^1(v - 1)}{\Phi_p^1(\bar{p}, 1, \lambda)}$$

Finally, using the elasticity of $\bar{p}$ with respect to λ, defined in (8) as $\varepsilon \equiv -\frac{\partial \bar{p}}{\partial \lambda}\frac{\lambda}{p} = -\frac{\lambda}{\Phi_p^1(\bar{p}, 1, \lambda)}$, where $0 < \varepsilon < 1$, this can be rewritten as:

$$p_2 - p_1 = \frac{\varepsilon}{\lambda}[\bar{p}(\lambda_1 - \lambda_2) - 2\Phi_v^1(v - 1)]$$

Since $-2\Phi_v^1 > 0$ for $v > 1$, the probability of detection will increase if the efficiency of enforcement rises in period 2 relative to period 1 (i.e., $\lambda_1 > \lambda_2$) and if the penalty under the VDP is set at a higher level.

Appendix V. Extension – Anticipated VDP over T periods

In the basic model of an anticipated VDP, taxpayers and the government optimize their behavior over two periods. This Appendix generalizes the results to a T-period model. Thereby, the VDP is offered in period 2 only and those opting into the VDP will be treated as compliant taxpayers in subsequent periods.

Taxpayers' choices in the first two periods remain the same as in the main text, but life goes on for another T-2 periods subsequently. The payoff from the three strategies (C,C), (E,E) and (E,V) reads as

$$U^c = Tu^c \tag{AV.1.a}$$
$$U_i^e(p) = Tu_i^e(\bar{p}) \tag{AV.1.b}$$
$$U^{vdp}(p_1) = u_i^e(p_1) + (1 - \tau v)y + (T - 2)u^c \tag{AV.1.c}$$

Where $\bar{p} = \frac{1}{T}\sum_i p_i$ is the average probability of detection over all periods. Equivalence of the payoffs implies the following threshold guilt levels:

$$g_l^T = \frac{\tau\left(\frac{T - 2 + v}{T - 1} - \overline{p_2}\varphi\right)}{1 - \overline{p_2}} \qquad g_h^T = \frac{\tau(\rho - p_1\varphi)}{1 - p_1} \tag{AV.2}$$

Where we now define, in slight abuse of notation, $\overline{p_2} = \frac{1}{T-1}\sum_{i=2}^{T} p_i$.

Note that g_l^T converges towards $\frac{\tau(1-\overline{p_2}\varphi)}{1-\overline{p_2}}$ as T tends to infinity, which is (approximately) the same threshold guilt level as for taxpayers deciding between straight compliance or evasion.

Revenue over T periods is

$$R = [T(1 - g_h^T) + (g_h^T - g_l^T)[v + (T - 1)] + [p_1 g_h^T + (T - 1)\overline{p_2}]\varphi - \sum_{i=1} C(p_i, g_i)]\tau y$$

The first order condition with respect to the VDP fine is, as before:

$$2(g_h^T - g_l^T) = \frac{\alpha\tau(\overline{p_2} - p_1)}{(1 - \overline{p_2})(1 - p_1)} \tag{AV.3}$$

Implying that the optimal penalty is given by

$$v^{**} = 1 + \frac{1}{T}\frac{\overline{p_2} - p_1}{1 - \bar{p}}\left(\varphi - 1 - \frac{\alpha}{2}\right) \tag{AV.4}$$

This provides several insights. First, considered over a long horizon a one-off anticipated VDP in period 2 remains incentive compatible only if evasion exerts a negative externality ($\alpha > 0$). Second, when technological progress leads to an increasing trajectory of future detection probabilities (rather than a one-time jump between period 1 and period 2), a larger share of individuals will opt into the VDP in period 2. This can be seen from noting that if $\overline{p_2} > p_2$, equation (A3) implies that $\frac{\partial(g_h^T - g_l^T)}{\partial \overline{p_2}} > 0$. Third, and relatedly, the revenue maximizing penalty in the VDP is declines as T gets larger. In fact, as T goes to infinity (and in the absence of discounting) the optimal VDP penalty converges to $v^{**} = 1$, i.e., no penalty is provided to previous evaders.

References

Allingham, M. G., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1, 323–338.

Alm, J., M. McKee and W. Beck. 1990. Amazing Grace: Tax Amnesties and Compliance. *National Tax Journal* 43, 23-37.

Alstadsæter, A., N. Johannesen and G. Zucman. 2018. Who Owns the Wealth in Tax Havens? Macro Evidence and Implications for Global Inequality. *Journal of Public Economics* 162: 89–100.

Alstadsæter, A., N. Johannesen, and G. Zucman. 2019. Tax Evasion and Inequality. *American Economic Review*, 109 (6): 2073-2103.

Andreoni, J. 1991. The desirability of a permanent tax amnesty. *Journal of Public Economics* 45, 143–159.

Baer K. and E. Le Borgne. 2008. Tax Amnesties: Theory, Trends, and Some Alternatives. International Monetary Fund, Washington DC.

Beer S., M. Coelho and S. Leduc. 2019. Hidden Treasure: The Impact of Automatic Exchange of Information on Cross-Border Tax Evasion. IMF Working Paper 19/286, Washington DC.

Benedek, D., M. Grote, G. Jackson, M. Markevych, L. Sofrona and and C. Waerzeggers. 2022. "Voluntary Disclosure Programs – Design, Principles, and Implementation Considerations", IMF Technical Note and Manual 2022/002, Washington DC.

Buchanan, John and Lucilla McLaughlin. 2017. "Revenue Administration: Implementing a High-Wealth Individual Compliance Program". Technical Notes and Manuals, April 2017, Fiscal Affairs Department, IMF.

Fischbacher, U., S. Gaechter and E. Fehr. 2001. Are people conditionally cooperative? Evidence from a public goods experiment. *Economic Letters* 71(3).

Franzoni, L.A. 1996. Punishment and Grace: On the Economics of Permanent Amnesties, mimeo.

Gupta S. M. Keen, A. Shah. G. Verdier. 2017. Digital Revolutions in Public Finance, International Monetary Fund, Washington DC.

International Monetary Fund, 2022, Fiscal Monitor Chapter 2: Coordinating Taxation Across Borders, Washington D.C.

Johannesen, N., Zucman, G. 2014. The end of bank secrecy? An evaluation of the G20 tax haven crackdown. *American Economic Journal: Economic Policy* 6, 65–91.

Keen, M., and J. Slemrod, 2017, Optimal Tax Administration. IMF Working Paper No. 17/8, Washington DC.

Langenmeyr. D. 2017. Voluntary Disclosure of Evaded Taxes – Increasing Revenues, or Increasing Incentives to Evade? *Journal of Public Economics* 151, 110-125.

Leenders, W., A. Lejour, S. Rabate and M. van 't Riet, 2023, Offshore Tax Evasion and Wealth Inequality: Evidence from a Tax Amnesty in the Netherlands, *Journal of Public Economics* 217.

Leonard, H.B. and R. Zeckhauser. 1987. Amnesty, Enforcement, and Tax Policy, in: Tax Policy and the Economy, Volume 1, 1987, pp 55-86, National Bureau of Economic Research.

Malik, A. S., Schwab, R. M., 1991. The economics of tax amnesties. *Journal of Public Economics* 46, 29–49.

OECD. 2015. Update on Voluntary Disclosure Programmes: A Pathway to Tax Compliance, Paris.

Slemrod, J. B., 2007. Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives* 21, 25–48.

Stella, P., 1991. An economic analysis of tax amnesties. *Journal of Public Economics* 46, 383–400.

Zucman, G. 2013. The missing wealth of nations: Are Europe and the U.S. net debtors or net creditors? *Quarterly Journal of Economics* 128, 1321–1364.

PUBLICATIONS