Constrained Efficient Fine-cum-Tax Rate Structures: The Case of Constant Relative Risk Aversion

By K. L. GLEN UENG† and C. C. YANG†,‡

†*National Cheng-Chi University, Taipei, Taiwan* ‡*Academia Sinica, Taipei, Taiwan*

Final version received 6 November 2003.

This paper explores the possibility that fines and tax rates impose different welfare costs on different types of taxpayer, so that the government may want to apply different fine/tax rate treatments to different types of taxpayer. Under the assumption that taxpayers exhibit constant relative risk aversion, we characterize the constrained Pareto-efficient fine/tax rate structure, showing that: (i) the co-existence of compliers (the above-ground economy) and evaders (the underground economy) is typical at the Pareto optimum, and (ii) Pareto efficiency requires compliers to face tax rates no higher and fines no lower than those faced by evaders.

INTRODUCTION

The tax evasion literature has become huge since the seminal work of Allingham and Sandmo (1972) and Yitzhaki (1974).[1] Of the many possible research directions, Border and Sobel (1987), Mookherjee and Png (1989), Cremer and Gahvari (1995), Schroyen (1997), and Chander and Wilde (1998), among others, have considered the optimal income taxation in the presence of tax evasion, allowing for the possibility that income can be verified only through costly audits. This line of study focuses on the properties of an optimal *audit* strategy as opposed to those of an optimal tax schedule. As for fines, they are either set as large as possible to provide the maximum incentive for truth-telling (the so-called 'Maximal Punishment Principle'),[2] or they are assumed to be exogenously predetermined.

In this paper, fines are neither set at their maximum nor given exogenously. We focus on the properties of an optimal *fine* schedule as opposed to those of an optimal *tax* schedule, emphasizing the possibility that fines and tax rates impose different welfare costs on different types of taxpayer, so that the government may want to allow different types of taxpayer different fine-cum-tax rate treatments. Most tax laws specify fines and tax rates explicitly, delegating the practice of auditing to the tax agency. In this sense, we address an issue that concerns the normative side of tax laws.

We consider a model in which taxpayers differ not only in terms of income but also in terms of detection probability if they choose to evade tax. The income dimension embodies a taxpayer's ability to earn income. As to the detection probability dimension, this measures how likely it is that evasion will be detected by the tax agency and hence may be regarded as reflecting a taxpayer's ability to conceal income (net of the tax agency's audit effort).[3] By employing such a model, we characterize constrained efficient fine/tax structures under three main assumptions.[4]

© The London School of Economics and Political Science 2004

First, we confine the study to the important class of utility functions that exhibit constant relative risk aversion (CRRA). An interesting property resulting from the imposition of CRRA is that taxpayer differences in compliance rate (i.e. the percentage of income reported to the tax agency) are all attributed to the differences in the probability of evasion detection. This property will be shown to be important, for it implies that the 'single crossing' of preferences across types is satisfied in our setting. Another interesting property implied by the CRRA assumption is that the identification of a taxpayer's income matters only if one considers the choice of a particular fine/tax rate structure within the set of efficient structures. As far as the characterization of efficient fine/tax rate structures is concerned, we can in effect ignore the income dimension of taxpayers and focus on the detection probability dimension. This clean separation seems to be a useful benchmark for further studies of the non-CRRA cases.

Second, we assume that income can vary continuously across taxpayers but that there are only two types of taxpayer in terms of the probability of evasion detection. Specifically, taxpayers are assumed to face either a high or a low probability of detection if they evade tax. The setup of two-class economies has been popular in the taxation literature and appears in Mirrlees (1975), Stiglitz (1987), Cremer and Gahvari (1995), Schroyen (1997) and many other studies. According to Mirrlees,

> The two-class economy is sufficiently rich in possibilities to be worth examining, for it captures in a crude way many of the two-way distinctions that strike us as interesting and relevant—rich and poor, capitalists and workers, urban and rural dwellers, etc. (Mirrlees 1975, pp. 27–8)

The list of two-class economies may be extended to include the case where there are compliers (the above-ground economy) and evaders (the underground economy). As will be seen, our two-type setup (taxpayers face either a high or a low probability of evasion detection) does result in this two-class economy.

Third, both the ability to earn and the ability to conceal income (hence both the income earned and the detection probability faced by a taxpayer in our setting) are assumed to be private information, unknown to the government. A taxpayer's income is typically assumed to be private information in the tax evasion literature. As for the assumption that detection probability is private information, this may be justified as follows. Tax evasion and its associated underground economies seem to exist everywhere as long as there are taxes.[5] Taxpayers are assumed to face either a high or low probability of evasion detection in our two-class economy. One may view those taxpayers who face a low detection probability as a proxy for individuals who work underground and evade tax. However, the government *a priori* does not know who is working in the underground economy. As a result, the government *a priori* cannot identify the low or the high detection probability associated with a particular taxpayer.[6]

Under the above three assumptions, we derive two interesting results: (i) the co-existence of compliers (the above-ground economy) and evaders (the underground economy) is typical at the Pareto optimum; and (ii) Pareto efficiency requires that compliers face tax rates no higher and fines no lower than evaders.

In his survey of the literature on optimal taxation, Stiglitz (1987) considers an economy in which taxpayers differ in terms of their ability to earn income and where this ability is private information. Following the tradition of Mirrlees (1971), Stiglitz assumes that income is perfectly and costlessly observable and there is therefore no tax evasion. He characterizes the constrained Pareto-efficient *tax* structures. By contrast, this paper considers, under the CRRA assumption, an economy in which taxpayers differ basically in terms of their ability to conceal income. This ability is private information and is reflected in different probabilities of evasion detection faced by the taxpayers. We characterize the constrained Pareto efficient *fine/tax rate* structures.

The rest of this paper is organized as follows. Section I introduces our model. Section II proves three useful lemmas, paving the way for the subsequent analysis. Section III is the main part of the paper, in which we identify and characterize the set of Pareto efficient fine/tax rate structures. Section IV concludes.

I. A Model of Tax Evasion

Most tax laws specify fines and tax rates explicitly, delegating the practice of auditing to the tax agency. Taking this statement seriously, one should model the problem as a two-stage game: the government (say, the US Congress) chooses the fine/tax rate structure in the first stage, while the tax agency (say, the US Internal Revenue Service) chooses the audit rule in the second stage. For tractability, however, we shall treat the tax agency's audit rule in the second stage of the game as a black box and summarize the outcomes resulting from the black box simply by the data that taxpayers face either a high or a low probability of detection if they choose to evade tax. For example, it is easier for the evasion of a wage/salary earner to be detected by the tax agency than that of someone who is self-employed.

Existing models on the formulation of a tax agency's audit policy can be divided into two groups: the principal–agent model (auditing with commitment) and the game-theoretic model (auditing without commitment). In their critical assessment of the existing models, Andreoni *et al*. (1998) point out that a crucial assumption maintained in these models is that taxpayers can correctly know or deduce the tax agency's audit rule. However, the evidence indicates otherwise: 'many taxpayers have only a rough idea of the average probability of audit in their class, and most have little idea as to how this probability changes with the level of income reported' (p. 833, fn. 42). Indeed, the so-called 'discriminant index function', a secret audit formula developed by the Internal Revenue Service to enforce tax compliance, is strictly guarded and unknown to the public.[7] In view of this evidence, our short-cut approach as described above may not be a bad research strategy. This 'partial equilibrium' or truncated model is certainly not fully satisfactory, but it is not inconsistent with the above evidence; at the same time, it may have the merit of serving as a stepping stone to the more complete model.

The tax agency as a rule makes its own audit rule and carries out its auditing independently. The government can of course control the tax agency's auditing indirectly via adjusting the size of the budget appropriated for the tax agency. One can imagine that the probabilities of evasion detection faced by taxpayers will be higher (lower) in general if the budget size appropriated for the tax agency becomes larger (smaller). Our derived results remain qualitatively robust with respect to these possible changes in the tax agency's budget. This is true because the probabilities of evasion detection faced by taxpayers in our model are *arbitrarily* given and hence can correspond to any positive budget size.

The model

Consider an economy in which taxpayers are characterized by two parameters: income level, $Y \in (0, \infty)$, and probability of evasion detection, $0 < p_L < p_H < 1$. Each taxpayer receives an exogenous income and faces either a high probability of evasion detection, p_H, or a low probability, p_L. The income dimension embodies a taxpayer's ability to earn income, while the detection probability dimension reflects his or her ability to conceal income (net of the tax agency's audit effort). Both parameters are private information, unknown to the government. Although the government cannot identify a taxpayer's income *a priori*, we assume that it knows the average income of both type-p_H and type-p_L taxpayers.

As noted by Chander and Wilde (1998), much of the tax evasion literature treats income as exogenous. The focus of this line of the literature is on the informational constraint arising from the fact that a taxpayer's income cannot be directly observed. We follow this focus here.

Except for their earning different levels of income and/or facing either a high or a low probability of evasion detection, all taxpayers are assumed to be identical. Let $U(.) : R_+ \to R_+$ be a smooth von Neumann–Morgenstern utility function with $U' > 0$ and $U'' < 0$. A taxpayer (characterized by p and Y) is assumed to declare income X ($\leqslant Y$) so as to maximize the expected utility:

$$EU(p, Y) = (1 - p)U(A) + pU(B), \tag{1}$$

where $A \equiv Y - tX$, $B \equiv Y - tX - Ft(Y - X)$, and $F > 1$ and $t < 1$ denote the proportional fine and tax rate, respectively. It is clear from (1) that, if $F \leqslant 1$, all taxpayers will declare zero income ($X = 0$) and government revenue collected will come only from those taxpayers who are detected by the tax agency. The case of $F \leqslant 1$ is rather unrealistic and uninteresting. We rule it out by assuming that the amount of revenue collected would be too low to meet the needs of the economy if $F \leqslant 1$. If $t \geqslant 1$, the tax rate imposed would be *de facto* a confiscation or punishment. To our knowledge, there is no single case in modern democratic societies in which such confiscation tax rates are imposed.

The setup of equation (1) assumes: (i) that true income is unknown to the tax authority but will be discovered once evasion gets detected,[8] and (ii) that the caught evader will be fined, and a penalty levied on the amount of evaded tax, as is the case under most tax laws. This setup has been popular in the tax evasion literature since the seminal work of Allingham and Sandmo (1972) and Yitzhaki (1974). Our tax evasion model differs from the Allingham–Sandmo–

Yitzhaki models on two interesting counts. First, we regard the probability of evasion detection, p, as a type dimension of taxpayers, allowing for taxpayer heterogeneity not only in terms of income but also in terms of the ability to conceal income. Second, as will be clear, both fines and tax rates imposed may be individual-specific, depending upon a taxpayer's type.

Given that $F > 1$, there will be no evasion (i.e. $X = Y$) for a taxpayer of type (p, Y) if $pF \geqslant 1$. On the other hand, tax evasion will take place (i.e. $X < Y$) for the taxpayer if $pF < 1$. Note that whether or not a taxpayer will engage in tax evasion is determined solely by detection probability p and fine F. The reason for this result is that the condition $pF = 1$ happens to represent a fair gamble for tax evasion in (1). It is well known that a risk averter takes no part of an unfavourable or barely fair gamble, but always takes some part of a favourable gamble (Arrow 1970, pp. 99–100).[9]

When tax evasion occurs, the first-order condition from the maximization of (1) is

$$p(F-1)U'(B) - (1-p)U'(A) = 0. \quad (2)$$

The second-order condition is satisfied under the assumption that $U'' < 0$. Let $X(F,t;\, p,Y)$ be the solution to (2). For convenience, we denote $X(F,t;\, p,Y)$ by the symbol X throughout the rest of the paper.

Taxpayer preferences over fines and tax rates

In this subsection we derive taxpayer preferences over fines and tax rates resulting from the above model.

Consider a taxpayer who is characterized by p and Y. Let F^M satisfy the equality $pF^M = 1$. The taxpayer will not evade tax if the fine levied is no less than F^M; otherwise, the taxpayer will evade tax. The indirect utility function from (1) will be

$$v(F, t;\; p, Y) = U[(1-t)Y] \quad (3)$$

if $F \geqslant F^M$, which obviously implies that

$$\left.\frac{\mathrm{d}t}{\mathrm{d}F}\right|_{v=\bar{v}} = 0. \quad (4)$$

That is, the marginal rate of substitution between fines and tax rates (MRS) is zero whenever evasion does not take place.

On the other hand, the indirect utility function from (1) and (2) will be

$$v(F, t; p, Y) = (1-p)U(A) + pU(B) \qquad \text{if } F < F^M, \quad (5)$$

which in turn implies that

$$-\left.\frac{\mathrm{d}t}{\mathrm{d}F}\right|_{v=\bar{v}} = \frac{t}{F}(1-x), \quad (6)$$

where $x = X/Y$, the compliance rate or the percentage of income that is actually reported to the tax authority. The derivation of (6) has made use of the first-order condition (2). This result indicates that the MRS is positive whenever evasion takes place (i.e. $x < 1$). Note that $x = 1$ if $F \geqslant F^M$, and hence (4) can be regarded as a special case of (6) with $x = 1$.

According to (6), there exists a substitution between fines and tax rates for the taxpayer if he or she evades tax (i.e. $x < 1$); but the substitution ceases to exist if there is no evasion (i.e. $x = 1$). When a taxpayer does not evade tax, only the tax rates count (see equation (3)). By contrast, when a taxpayer evades tax, higher tax rates cause a decrease in utility and so do higher fines. This implies a substitution between fines and tax rates for the taxpayer whenever evasion occurs. Figure 1 depicts indifference curves for a typical taxpayer. Note that the indifference curve displays no kink at the point $F = F^M$. It is obvious that, the higher the indifference curve, the lower will be the utility of the taxpayer.

Equation (6) gives a nice characterization for taxpayer preferences over fines and tax rates and can be rewritten as

$$(7) \qquad -\left.\frac{\mathrm{d}\ln t}{\mathrm{d}\ln F}\right|_{V=\bar{V}} = 1 - x.$$

This result tells us that by keeping utility constant a taxpayer's elasticity of tax rates with respect to fines is completely determined by his or her compliance rate; moreover, the higher the compliance rate, the lower the elasticity. Thus, taxpayers with different compliance rates in a sense have different degrees of tolerance towards fines relative to tax rates. Those taxpayers who evade completely will have the lowest tolerance (with the elasticity equal to 1), while those taxpayers who comply completely will have the highest tolerance (with the elasticity equal to 0). The focus of this paper is to explore an implication following from this result: fines and tax rates impose different welfare costs on

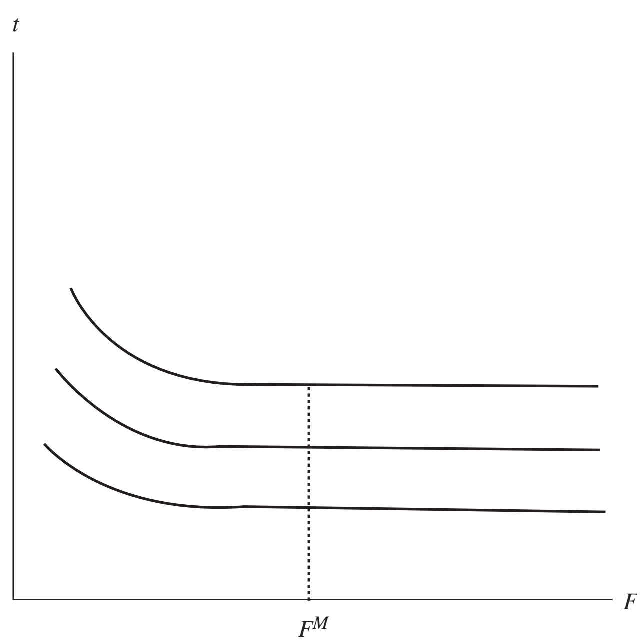

FIGURE 1. Indifference curves for a typical taxpayer

different types of taxpayer, and hence the government may want to afford different fine/tax rate treatments to different types of taxpayer. We confine our exploration to the important class of utility functions that exhibit constant relative risk aversion.

Constant relative risk aversion

While according to (2) the *absolute* amount of income concealed is predicted to rise with income if absolute risk aversion is decreasing, the *proportion* of income concealed may rise, decline or remain unchanged. To make a definite prediction, we need to restrict taxpayer preferences further with respect to relative risk aversion.

The assumption of decreasing absolute risk aversion seems to accord with both intuition and everyday observation (see Arrow 1970, chapter 3). It is, however, less obvious empirically whether relative risk aversion is decreasing, increasing or constant. As a first-order approximation, and for analytical tractability, it is fairly standard in a variety of areas to assume that relative risk aversion is constant. We conform to this standard practice and confine our analysis to the class of utility functions that satisfy the following.

Assumption 1. Taxpayer preferences exhibit constant relative risk aversion (CRRA).

Heterogeneity between taxpayers in our model ultimately stems from the different characteristics of taxpayers, that is from differences in either the detection probability they face or the income level they earn (or both). It is known that the compliance rate will rise/decline/remain unchanged with income depending on whether the relative risk aversion is, respectively, increasing/decreasing/constant (see Cowell 1985). With Assumption 1, the compliance rate is independent of the income level; that is,

$$x_Y \equiv \frac{\partial x}{\partial Y} = 0. \tag{8}$$

As a consequence of (8), the different compliance rates under the CRRA assumption are all attributed to differences in the probability of evasion detection that taxpayers face.

Note that $U(A) = U[Y \cdot (1 - tx)]$ and $U(B) = U\{Y \cdot [1 - tx - Ft(1 - x)]\}$. With setup (1), the CRRA assumption will induce an expected utility function where p and Y appear separately. For example, consider the utility function $U(c) = \ln c$. Substituting this function into (1) yields

$$EU(p, Y) = \ln Y + \{(1 - p)\ln(1 - tx) + p\ln[1 - tx - Ft(1 - x)]\}, \tag{1'}$$

in which the compliance rate x is independent of Y according to (8). It is clear that (1′) is decomposed into two separate parts, one involving the parameter p and the other involving the parameter Y. Since the latter part is only a constant term and appears separately, our two-parameter characterization (p, Y) of the taxpayers *de facto* reduces to one parameter (p) as far as characterizing efficient fine/tax rate structures is concerned. This property holds for other CRRA utility functions as well.

Finally, it is important to observe that, with the imposition of the CRRA assumption, the MRS (6) is independent of Y and strictly decreasing in p as long as $x<1$. Thus, under the CRRA assumption, the 'single crossing' of preferences across types is satisfied in our framework if we restrict our attention to the evasion case.

II. Three Useful Lemmas

Before proceeding to the main results, we present three useful lemmas in this section.

Given a fine/tax rate (F, t), the expected revenue collected from a taxpayer (characterized by p and Y) equals

$$r(F, t; p, Y) = Y \cdot \{(1-p)tx + p[tx + Ft(1-x)]\}. \tag{9}$$

The revenue-neutral transformation between fines and tax rates (*RNT*) for a taxpayer equals

$$-\left.\frac{\mathrm{d}t}{\mathrm{d}F}\right|_{r=\bar{r}} = \frac{r_F}{r_t}, \tag{10}$$

where $r_F \equiv \partial r/\partial F$ and $r_t \equiv \partial r/\partial t$. It can be shown that the sign of (10) is non-negative. If the taxpayer does not evade with $pF \geqslant 1$, then the sign of (10) will become zero since $r_F = 0$.

We are now ready for the following three lemmas. All the proofs are relegated to the Appendix.

Lemma 1. Raising fines along taxpayer *i*'s indifference curve (with the tax rates reduced if necessary), the expected revenue collected from taxpayer *j* will increase/decrease/remain unchanged depending on whether taxpayer *i*'s MRS is, respectively, smaller than/greater than/equal to taxpayer *j*'s corresponding RNT.

Lemma 2. From Assumption 1 (CRRA), the MRS (6) is smaller than (equal to) the RNT (10) at any fine/tax rate for a taxpayer if evasion occurs (does not occur).

Lemma 3. From Assumption 1 (CRRA), raising fines along a taxpayer's indifference curve increases (leaves unchanged) the expected revenue collected from this taxpayer if he or she evades tax (does not evade tax).

Given a fine/tax rate (F, t), one can rewrite (1) as

$$EU(p, Y) = U(Y - r(p, Y) - \theta(p, Y)), \tag{1''}$$

where $\theta(p, Y) > 0$ is the risk premium that a risk-averse taxpayer (characterized by p and Y) would be ready to pay in order to eliminate the exposure to audit risk. Yitzhaki (1987) calls this risk premium the 'excess burden of tax evasion'. It is an excess burden because $\theta(p, Y)$ represents a deadweight loss beyond what would be imposed if $r(p, Y)$ were somehow collected by a lump-sum tax. On the basis of (1″), Lemma 3 then implies that raising fines along a taxpayer's indifference curve lessens (leaves unchanged) the excess burden of tax evasion imposed upon this taxpayer if he or she evades tax (does not evade tax).

III. Characterizing Efficient Fine/Tax Rate Structures

In this section we characterize constrained efficient fine/tax rate structures. We first derive the Pareto-efficient programme in question.

Deriving the Pareto-efficient programme

The first step in the derivation is to show that, owing to the CRRA assumption, we can disregard all incentive compatibility (IC) constraints associated with the income dimension of taxpayers.

As far as the detection probability dimension is concerned, there are two types of taxpayer in our economy: those with low and high probability of evasion detection, i.e. p_H and p_L, with $0<p_L<p_H<1$. Without loss of generality, we assume that the two types of taxpayer are of equal size. Let (F_i, t_i) be the fine/tax rate intended for taxpayers of type p_i $(i = H, L)$. The IC constraints facing the government are:

$$v(F_H, t_H; p_H, Y) \geqslant v(F_L, t_L; p_H, Y) \quad \forall Y;$$

$$v(F_L, t_L; p_L, Y) \geqslant v(F_H, t_H; p_L, Y) \quad \forall Y.$$

Note that the IC constraints above need to hold for all possible Y associated with type-p_H and type-p_L taxpayers. This means that the government may face an infinite number of IC constraints, since income can vary continuously across taxpayers in our model. Fortunately, by invoking Assumption 1 (CRRA), the indirect utility function $v(\cdot)$ can be expressed as[10]

$$v(F, t; p, Y) = V(F, t; p) \cdot g(Y).$$

Using this functional separation between p and Y, the possibly infinite number of IC constraints can be reduced to the following two:

$$V(F_H, t_H; p_H) \geqslant V(F_L, t_L; p_H);$$

$$V(F_L, t_L; p_L) \geqslant V(F_H, t_H; p_L).$$

That is, the same IC constraint applies to the same type-p taxpayers regardless of their incomes. This property *de facto* reduces our two-dimensional screening problem to a one-dimensional one, making the analysis tractable.

The Pareto-efficient programme, which the government seeks to solve, is[11]

$$(\alpha) \qquad \max_{\{F_i, t_i\}} V(F_L, t_L; p_L)$$

$$\text{s.t. } V(F_H, t_H; p_H) \geqslant \bar{V}_H$$

$$V(F_H, t_H; p_H) \geqslant V(F_L, t_L; p_H)$$

$$V(F_L, t_L; p_L) \geqslant V(F_H, t_H; p_L)$$

$$r_H + r_L \geqslant \bar{R},$$

where $\bar{V}_H$ is the feasible utility level given for type-p_H taxpayers and $\bar{R}>0$ denotes the exogenously required government revenue. The second and the third inequality above are the incentive compatibility (IC) constraint, while the last one is the revenue constraint. Since the government is assumed to know the average income of both type-p_H and type-p_L taxpayers, it can calculate the revenue collected from the economy (i.e. $r_H + r_L$) without identifying a particular taxpayer's income. It is clear that the set of efficient fine/tax rate structures is associated with the solutions to varying $\bar{V}_H$ in programme (α).

The first-best or perfect-information solution

If information were perfect, the government could ignore the IC constraint in programme (α), and hence the solution would require that

$$MRS_H = RNT_H \quad \text{and} \quad MRS_L = RNT_L.$$

That is, MRS (6) is set to equal RNT (10) for each type of taxpayer.

Define $F_i^M \equiv 1/p_i$ ($i = H, L$). According to Lemma 2, the above first-best solution implies that the government should set fines at the maximum with $F_H = F_L = F_L^M$[12] and collect the required revenue by imposing tax rates only. This solution obviously eliminates the 'excess burden of tax evasion' as defined by Yitzhaki (1987).

Let us denote the highest feasible $\bar{V}_H$ in our model by $\bar{V}_H^M$. As will be seen later, unless the $\bar{V}_H$ given in programme (α) happens to equal $\bar{V}_H^M$, the first-best solution will violate the IC constraint. In what follows we characterize first the constrained efficient fine/tax rate solutions associated with $\bar{V}_H < \bar{V}_H^M$, and then the solution associated with $\bar{V}_H = \bar{V}_H^M$.

Constrained efficient fine/tax rate structures when $\bar{V}_H < \bar{V}_H^M$

We derive several properties resulting from the Pareto-efficient programme under the CRRA assumption. We then summarize these properties with a proposition. All the proofs are relegated to the Appendix. Following the convention in the principal–agent model, type-p_H taxpayers are assumed to choose (F_H, t_H) rather than (F_L, t_L) if they are indifferent between them.

Property 1. Pareto efficiency requires that the revenue constraint be binding, i.e. that $r_H + r_L = \bar{R}$.

Property 2. Pareto efficiency requires that the IC constraint for type-p_H taxpayers be binding, i.e. that $V(F_H, t_H; p_H) = V(F_L, t_L; p_H)$.

Property 3. $F_H \geqslant F_H^M$ at the Pareto optimum.

Property 4. If $F_L < F_H^M$, the Pareto-efficient fine/tax rate structure will be separating as shown in Figure 2.

Property 5. If $F_L \geqslant F_H^M$, the Pareto efficient fine/tax rate structure will be pooling with $(F_H, t_H) = (F_L, t_L)$ as shown in Figure 3.

To sum up, we have the following proposition.

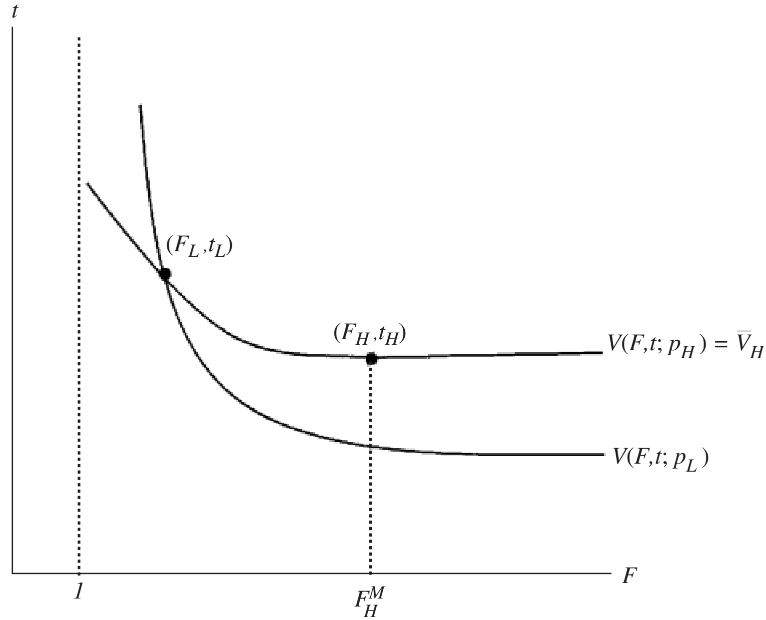

FIGURE 2. A separating fine-cum-tax rate structure

Proposition 1. From Assumption 1 (CRRA) and $\bar{V}_H < \bar{V}_H^M$, the Pareto-efficient fine/tax rate structure in a two-class economy with $0 < p_L < p_H < 1$ is either separating, as shown in Figure 2, or pooling, as shown in Figure 3.

According to Property 3, the revenue collected from type-p_H taxpayers (i.e. r_H) is fixed for each $V(F, t; p_H) = \bar{V}_H$ at the Pareto optimum. On the basis of Property 2, we then see that the highest $V(F_L, t_L; p_L)$ for each feasible $\bar{V}_H$ in programme (α) is to move (F_L, t_L) northwest along the indifference curve $V(F, t; p_H) = \bar{V}_H$ as far as possible and, at the same time, without violating the revenue constraint. If $F_L < F_H^M$ holds in the end, the separating case shown in Figure 2 will result. On the other hand, if $F_L \geqslant F_H^M$ holds in the end, the pooling case shown in Figure 3 will result.

In either case (separating or pooling), at the Pareto optimum (i) taxpayers of type p_H will comply, since $F_H \geqslant F_H^M$, and (ii) taxpayers of type p_L will evade tax, since $F_L < F_L^M$.[13] This leads to the following result.

Corollary 1. From Assumption 1 (CRRA) and $\bar{V}_H < \bar{V}_H^M$, Pareto efficiency in a two-class economy with $0 < p_L < p_H < 1$ implies the co-existence of compliers (the above-ground economy) and evaders (the underground economy).

A type-p taxpayer will evade tax if $pF < 1$. With $F_H \geqslant F_H^M$ at the Pareto optimum (i.e. Property 3), type-p_H taxpayers will not engage in tax evasion. As for type-p_L taxpayers, they will engage in tax evasion at the Pareto optimum if $\bar{V}_H < \bar{V}_H^M$. When interpreting this result, it is best to think of a society consisting of underground and above-ground economies. Taxpayers in the underground economy evade tax, while those in the above-ground economy do

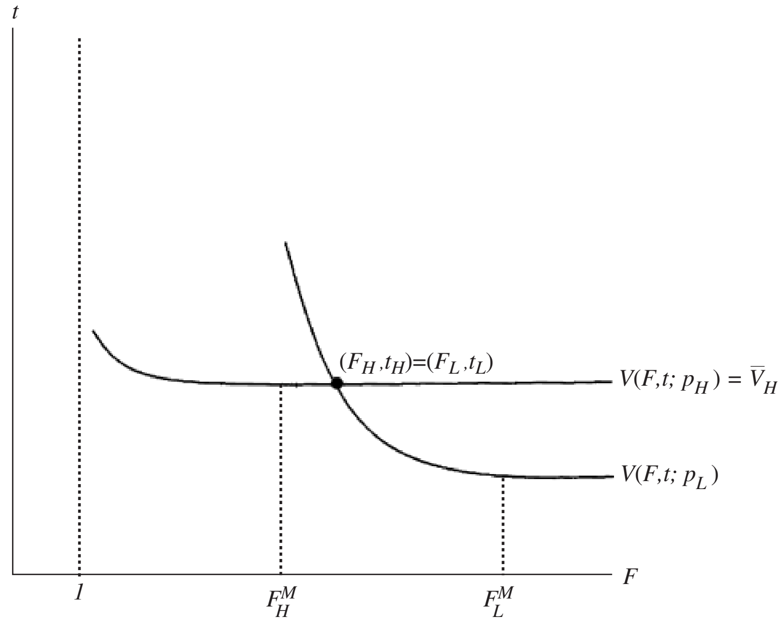

FIGURE 3. A pooling fine-cum-tax rate structure

not. A Pareto-efficient fine/tax rate structure allows for the co-existence of evaders and non-evaders, and it does not eliminate the underground economy.

Both Weiss (1976) and Stiglitz (1982) have suggested that it might be optimal for the government to introduce opportunities for income tax evasion, and that evasion might therefore be part of an optimal tax system. This argument, as noted by Slemrod and Yitzhaki (2002), depends on the second-best nature of the problem, which could mean that tax evasion might lessen the distortionary effect of income taxation associated with the labour–leisure choice. Our model does not allow for the labour–leisure choice, and so our reason for regarding the underground economy as desirable clearly differs from that of Weiss (1976) and Stiglitz (1982). Baldry (1984) has shown that a complete elimination of the underground economy is not efficient. His argument is based on the observation that a slight relaxation of complete enforcement will have little effect on the expected revenue collected, but will reduce the enforcement cost. Since the tax authority's enforcement efforts are assumed exogenous in our model, it is clear that our reason for regarding the underground economy as desirable also differs from that of Baldry (1984).[14]

In either case (separating or pooling), type-p_H taxpayers (compliers) face tax rates no higher and fines no lower than type-p_L taxpayers (evaders). Thus, our proposition also implies the following.

Corollary 2. From Assumption 1 (CRRA) and $\bar{V}_H < \bar{V}_H^M$, Pareto efficiency in a two-class economy with $0 < p_L < p_H < 1$ requires that taxpayers who comply face tax rates no higher and fines no lower than those who do not.

Pareto efficiency requires that those in the underground economy face a higher tax rate but a lower fine than those in the above-ground economy if separating occurs. If the pooling occurs, however, taxpayers in both the underground and above-ground economies face the same fine/tax rate at the Pareto optimum.[15]

Corollary 2 can be better understood by exploiting the MRS (6) and the IC constraint directly. Suppose that the bundle (F_H, t_H) depicted in Figure 4 belongs to the Pareto-efficient fine/tax rate structure. According to (6), and under Assumption 1, the slope of the indifference curve for type-p_L taxpayers will be steeper than that for type-p_H taxpayers at (F_H, t_H), and consequently the configuration of the indifference curves for the two types of taxpayer will look like that shown in Figure 4. To satisfy the IC constraint, it is clear that the bundle (F_L, t_L) must lie in the shaded region of Figure 4. This leads to the result that $t_H \leqslant t_L$ and $F_H \geqslant F_L$.

Another possible explanation for Corollary 2 is as follows. From (7), we see that taxpayers with different compliance rates in a sense have different degrees of tolerance towards fines relative to tax rates. Those taxpayers who evade completely will have the lowest tolerance (with the elasticity equal to 1), while those who comply completely will have the highest tolerance (with the elasticity equal to 0). At the Pareto optimum, taxpayers having more tolerance towards fines relative to tax rates (compliers) face higher fines but lower tax rates, whereas those having less tolerance (evaders) face lower fines but higher tax rates. Put differently, fines and tax rates impose different welfare costs on compliers and evaders, and so one may want to provide different treatments for them. From the viewpoint of fines alone, Corollary 2 may appear counterintuitive and contradictory to the principle of 'making the fine fit the crime'. However, the contradiction disappears to a large extent if one views the fine/tax rate structure as a whole.[16]

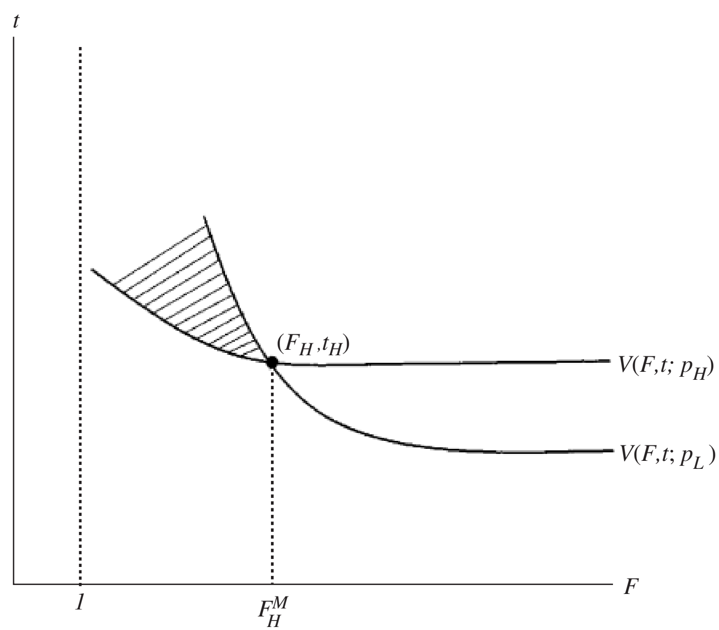

FIGURE 4. Compliers face tax rates no higher and fines no lower than evaders

The maximum-fine solution associated with $\bar{V}_H = \bar{V}_H^M$

Consider the pooling in Figure 5. The distinct feature of this pooling is that fines are set at their maximum (i.e. $F_H = F_L = F_L^M$) so that neither type of taxpayer evades tax. Let this maximum-fine pooling yield an amount of tax revenue exactly equal to the revenue requirement $\bar{R}$. Then it clearly belongs to the set of efficient fine/tax rate structures.

Since imposing fines higher than F_L^M will not increase revenue, while imposing tax rates lower than the maximum-fine pooling will surely decrease revenue, the tax rate $t_H = t_L$ shown in Figure 5 represents the lowest feasible tax rate compatible with the revenue requirement $\bar{R}$ as $F_H = F_L \geqslant F_L^M$. Type-$p_H$ taxpayers always comply in equilibrium according to Property 3. The lowest feasible tax rate thus implies the highest feasible utility reached by type-p_H taxpayers. This means that the $\bar{V}_H$ shown in Figure 5 is precisely the highest feasible utility $\bar{V}_H^M$ reached by type-p_H taxpayers in the Pareto-efficient programme.

With $\bar{V}_H = \bar{V}_H^M$, let us denote its corresponding minimal tax rate by t^m. All the configurations of pooling with $\bar{V}_H < \bar{V}_H^M$ must look like that shown in Figure 3 (in which type-p_L taxpayers evade tax since $F_L < F_L^M$) rather than that shown in Figure 5 (in which type-p_L taxpayers comply since $F_L = F_L^M$). The reason for this result is simple: a pooling with $\bar{V}_H < \bar{V}_H^M$ but $F_H = F_L \geqslant F_L^M$ must have $t_H = t_L > t^m$, and hence will yield an amount of revenue more than the revenue requirement $\bar{R}$. Based on Property 1, this more-than-required amount of revenue could then be used to raise the welfare of type-p_L taxpayers

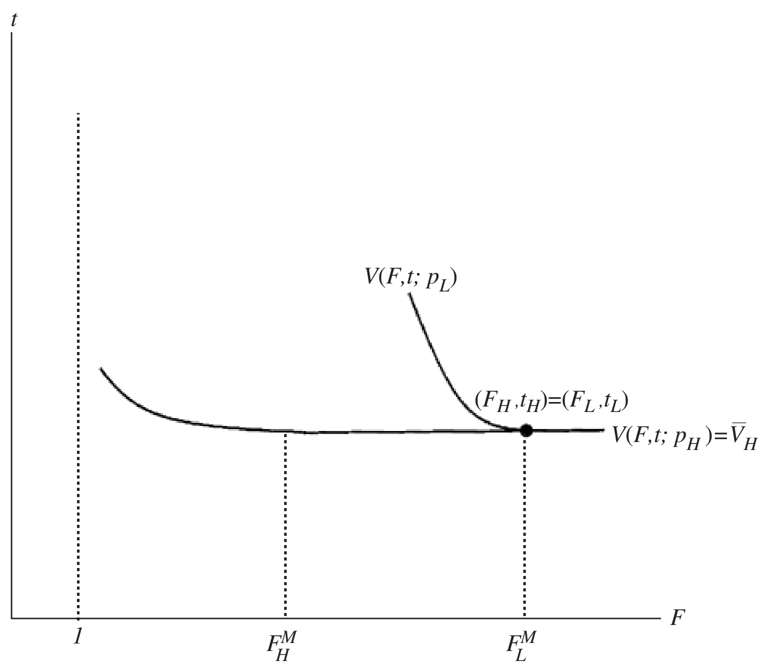

FIGURE 5. The maximum-fine solution

by reducing the fines intended for them.[17] This leads to the pooling with $F_H = F_L < F_L^M$, as shown in Figure 3.

Since type-p_H taxpayers always comply in equilibrium regardless of separating or pooling, the lower the tax rate they face, the higher the utility they will reach. When $\bar{V}_H = \bar{V}_H^M$, the maximum-fine solution $F_H = F_L = F_L^M$ will induce the lowest tax rate feasible, and hence the highest utility feasible, for type-p_H taxpayers.

Note that $V(F, t; p_H)$ and $V(F, t; p_L)$ can be compared directly through the horizontal portion of their respective indifference curves. Consider the utility function $U(c) = \ln c$, for example. The indifference curves corresponding to the horizontal portion of $V(F, t; p) = \bar{V}$ (i.e. the no-evasion portion of the indifference curve) are all represented by the same utility function $\ln(1 - t)$, regardless of whether taxpayers face a low or a high probability of evasion detection. Using this property, and remembering that the lower the indifference curve the higher the utility, we see from Figures 2 and 3 that the strict inequality $V(F_L, t_L; p_L) > V(F_H, t_H; p_H)$ will hold at the Pareto optimum when $\bar{V}_H < \bar{V}_H^M$. The equality $V(F_L, t_L; p_L) = V(F_H, t_H; p_H)$ will hold at the Pareto optimum if and only if $\bar{V}_H = \bar{V}_H^M$. (This is the case shown in Figure 5.[18])

When $\bar{V}_H = \bar{V}_H^M$, the maximum-fine solution $F_H = F_L = F_L^M$, together with $t_H = t_L = t^m$, happens to be incentive-compatible. However, when $\bar{V}_H < \bar{V}_H^M$ the maximum-fine solution will be associated with $t_H > t_L$. This fine/tax rate structure is not incentive-compatible. In this case, the IC constraint requires that $F_L < F_L^M$ in equilibrium. That is, type-p_L taxpayers are distorted to be evading tax rather than as they would behave in a first-best situation (i.e. $MRS_L = RNT_L$, so that there would be no 'excess burden of tax evasion'). This is an intuition behind our Corollary 1.[19]

Note that the pooling with $\bar{V}_H = \bar{V}_H^M$ is of measure zero in terms of all feasible values of $\bar{V}_H$. In this sense, the maximum-fine solution and its resulting elimination of the underground economy is atypical in the set of efficient fine/tax rate structures.

Remark. Given a pair of p_H and p_L, there will be a corresponding set of efficient fine/tax rate structures. This efficient set is composed of the solutions resulting from programme (α) by varying $\bar{V}_H$. It is important to recognize that there is no distinction between these solutions in the set as far as efficiency is concerned. This is true even for the maximum-fine solution. One might argue that the maximum-fine solution would be associated with the least enforcement cost and hence would be the most efficient solution in the set if enforcement costs were taken into account as well.[20] However, this argument is not valid in our framework. In our model the tax agency's audit rule is treated as a black box and the outcomes resulting from the black box are summarized by the data p_H and p_L. Within this setting, it seems only natural to think of the data p_H and p_L that result from the black box as corresponding to some fixed budget size for the tax agency. Indeed, one would expect that the data p_H and p_L that result from the black box would be different if there were a change in the appropriation of the tax agency's budget. Since a pair of p_H and p_L implies a corresponding fixed level of the tax agency's budget, it is clear that choosing the maximum-fine solution in the efficient set will not save the enforcement costs at all in our model.

IV. CONCLUDING REMARKS

This paper has characterized the efficient fine/tax rate structures under the CRRA assumption. By varying the value of $\bar{V}_H$ in programme (α), we can trace out the set of efficient fine/tax rate structures. This set is likely to be large. In fact, since there is an infinite number of feasible values of $\bar{V}_H$, the possible Pareto-efficient fine/tax rate structures are also infinite. It is obvious that the Pareto criterion alone cannot give us answers with regard to choices among these alternative efficient structures. To make a choice, interpersonal comparison or other considerations (say, the income dimension of taxpayers) are necessary.[21]

Although falling short of pinning down a particular fine/tax rate structure, our focus on identifying and characterizing the set of Pareto-efficient structures is consistent with the heart of the New New Welfare Economics (NNWE): the insistence on the distinction between efficiency and other considerations.[22] An important focus of NNWE is on the fact that economists could make important contributions to policy debates by presenting policy-makers with the Pareto-efficient set.

Our paper confines its study to the class of utility functions that exhibit constant relative risk aversion (CRRA). Like other models that rely upon the CRRA assumption, the usefulness of our results depends on the empirical relevance of this assumption.[23] Under this assumption, our two-parameter characterization (p, Y) of the taxpayers can be decomposed into two separate parts: the probability p dimension, and the income Y dimension. This clean separation may well serve as a useful benchmark for further studies, even if the CRRA assumption does not hold in the real world.

APPENDIX

Proof of Lemma 1

Denote taxpayer i's marginal rate of substitution by MRS_i and taxpayer j's revenue-neutral transformation by RNT_j. From (9),

(A1) $$\mathrm{d}r_j = (r_j)_F \mathrm{d}F + (r_j)_t \mathrm{d}t,$$

where r_j is the expected revenue collected from taxpayer j. Using (A1), we have

(A2) $$\frac{\mathrm{d}r_j}{\mathrm{d}F} = (r_j)_t \cdot \left(RNT_j + \frac{\mathrm{d}t}{\mathrm{d}F} \right).$$

Raising fines along taxpayer i's indifference curve then leads to

(A3) $$\frac{\mathrm{d}r_j}{\mathrm{d}F} = (r_j)_t \cdot (RNT_j - MRS_i).$$

Lemma 1 immediately follows from (A3). □

Proof of Lemma 2

We have $r_F = 0$ under the setup of (1) if $x = 1$. As a result, $MRS\,(6) = RNT\,(10)$ in this case.

Let $U(c) = c^{1-(1/\rho)}$, where $\rho > 0$ and $\rho \neq 1$. Then, with evasion,

(A4) $$X = \frac{(1 - D)}{1 - tD} \cdot Y,$$

where

$$D = \frac{1 - [(1-p)/(pF-p)]^{-\rho}}{tF}.$$

Denote the numerator of the D term by Δ. Note that $\Delta < 1, /\Delta_F < 0$, and

$$\text{(A5)} \qquad 1 - x = \frac{(1-t)D}{1-tD} = \frac{(1-t)\Delta}{t(F-\Delta)}.$$

We can also calculate

$$\text{(A6)} \qquad r_t = \frac{F(1-p\Delta)}{(F-\Delta)} \cdot Y,$$

$$\text{(A7)} \qquad r_F = \left[\frac{(1-pF)(1-t)(\Delta - F\Delta_F)}{(F-\Delta)^2} + \frac{p(1-t)\Delta}{F-\Delta} \right] \cdot Y.$$

From (A5)–(A7), it can be checked that the RNT (10) > the MRS (6) at any fine/tax rate if $x < 1$. This result is applicable to the case $U(c) = \ln c$ as well. □

Proof of Lemma 3

Applying Lemma 1 with $i = j$, Lemma 3 is then a consequence of Lemma 2. □

Proof of Property 1

There are two cases to consider: $F_L \geqslant F_L^M$ and $F_L < F_L^M$. The case where $F_L \geqslant F_L^M$ is associated with $\bar{V}_H = \bar{V}_H^M$ in programme (α). This maximum-fine case will be discussed later; here we focus on the case where $F_L < F_L^M$.

When $F_L < F_L^M$, we have at any fine/tax rate (i) $MRS_L > MRS_H$ by the 'single crossing' of preferences across types, and (ii) $RNT_L > MRS_L$ by Lemma 2. Putting (i) and (ii) together implies that $RNT_L > MRS_H$ at any fine/tax rate. Applying Lemma 1 then leads to the result that lowering F_L along an indifference curve of type-p_H taxpayers will decrease the expected revenue collected from type-p_L taxpayers. Note that lowering F_L along an indifference curve of type-p_H taxpayers will increase the utility of type-p_L taxpayers.

Now suppose that the fine/tax rate structure $\{(F_H, t_H), (F_L, t_L)\}$ is a Pareto optimum but that its associated revenue constraint is not binding (i.e. $r_H + r_L > \bar{R}$). Fix (F_H, t_H) and lower F_L along the indifference curve $V(F, t; p_H) = V(F_H, t_H; p_H)$ until the revenue constraint is binding. This move is feasible, since lowering F_L along the indifference curve $V(F, t; p_H) = V(F_H, t_H; p_H)$ will decrease the expected revenue collected from type-p_L taxpayers. Let (F'_L, t'_L) be the new fine/tax rate intended for type-p_L taxpayers. Note that (i) the new fine/tax rate structure $\{(F_H, t_H), (F'_L, t'_L)\}$ satisfies the IC constraint for type-p_H taxpayers since we lower F_L along the indifference curve $V(F, t; p_H) = V(F_H, t_H; p_H)$, so that $V(F_H, t_H; p_H) = V(F'_L, t'_L; p_H)$; and (ii) the new fine/tax rate structure $\{(F_H, t_H), (F'_L, t'_L)\}$ also satisfies the IC constraint for type-p_L taxpayers, since $V(F_L, t_L; p_L) \geqslant V(F_H, t_H; p_L)$ by assumption and $V(F'_L, t'_L; p_L) \geqslant V(F_L, t_L; p_L)$. However, lowering F_L along the indifference curve $V(F, t; p_H) = V(F_H, t_H; p_H)$ will increase the utility of type-p_L taxpayers and hence $V(F'_L, t'_L; p_L) > V(F_L, t_L; p_L)$. This contradicts the presupposition that the fine/tax rate structure $\{(F_H, t_H), (F_L, t_L)\}$ is a Pareto optimum. We conclude that the government revenue constraint must be binding at the Pareto optimum. □

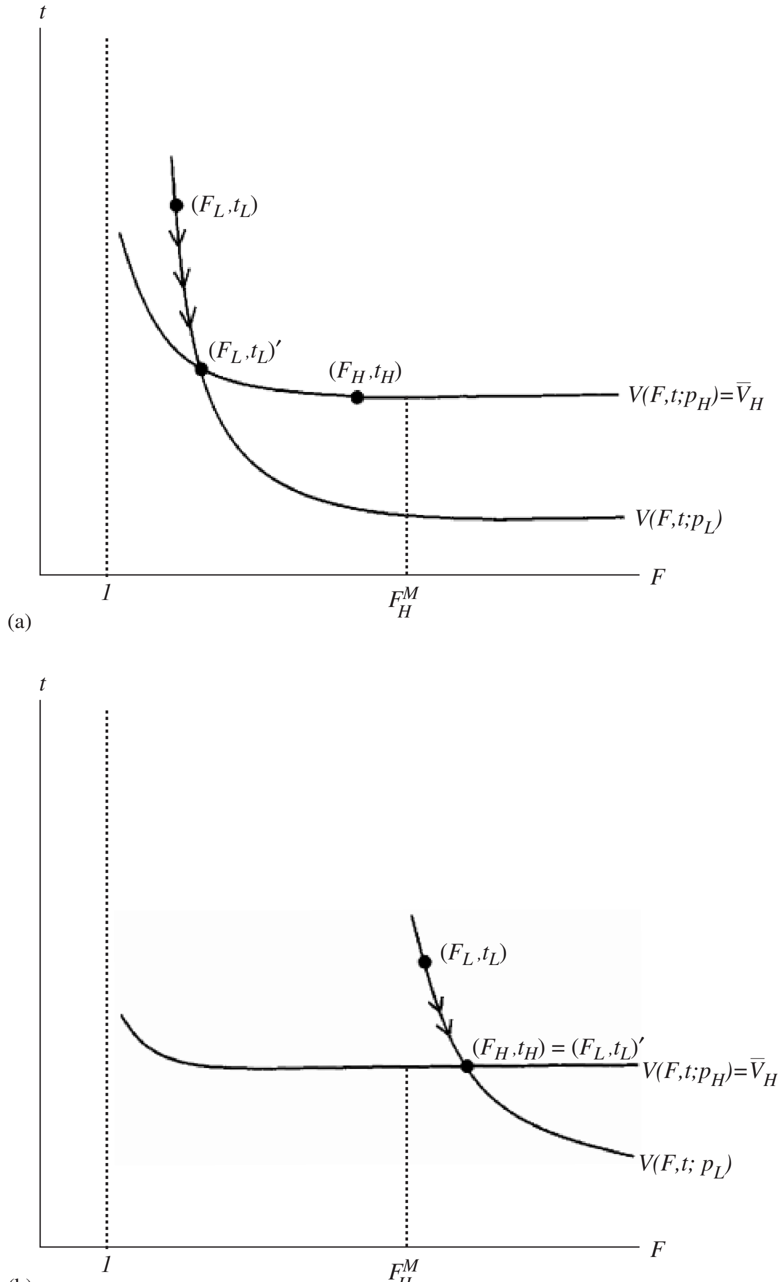

FIGURE A1. $V(F_H, t_H; p_H) = V(F_L, t_L; p_H)$ at the Pareto optimum

Proof of Property 2

Suppose that the fine/tax rate structure $\{(F_H, t_H), (F_L, t_L)\}$ is a Pareto optimum but that $V(F_H, t_H; p_H) > V(F_L, t_L; p_H)$, as shown in Figure A1(a) or (b) (the two possible cases). Then fix (F_H, t_H) but move (F_L, t_L) towards $(F_L, t_L)'$ along the indifference curve of type-p_L taxpayers, as shown in Figure A1. This move would obviously satisfy all the IC

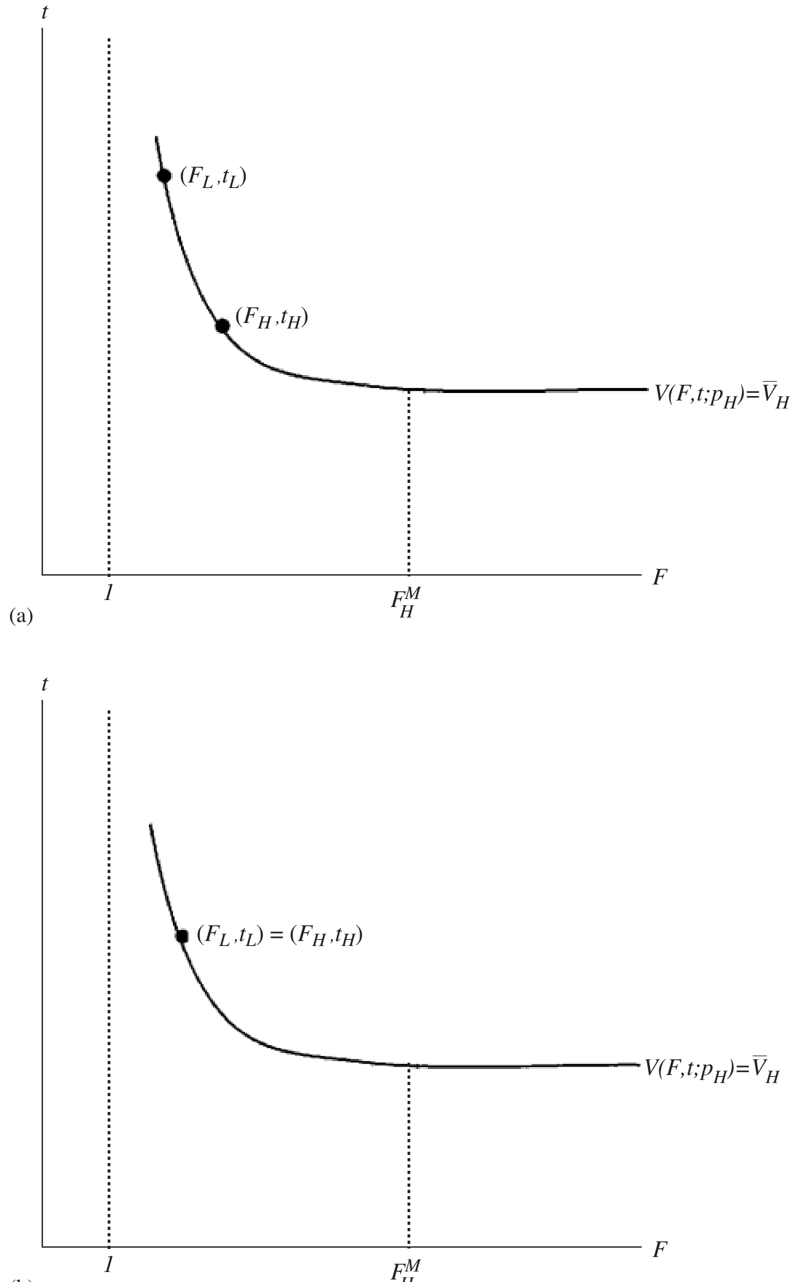

FIGURE A2. $F_H \geqslant F_H^M$ at the Pareto optimum

constraints and cause no taxpayers to become worse off. At the same time, it would result in an increase in revenue collected from type-p_L taxpayers and hence in total revenue, according to Lemma 3. This is in contradiction to Property 1 at the Pareto optimum. We therefore conclude that the strict inequality, $V(F_H, t_H; p_H) > V(F_L, t_L; p_H)$, cannot hold at the Pareto optimum. □

Proof of Property 3

On the basis of Property 2 and the IC constraint for type-p_L taxpayers, the fine/tax rate structure must look like that shown in Figures A2(a) or (b) (the two possible cases). Suppose that $F_H < F_H^M$ at the Pareto optimum. Then fix (F_L, t_L) but lower F_H towards F_H^M along the indifference curve $V(F, t; p_H) = \bar{V}_H$. This move would cause no taxpayers to become worse off and at the same time would satisfy all the IC constraints. However, the move would increase revenue collected from type-p_H taxpayers and hence the total revenue according to Lemma 3. This is in contradiction to Property 1 at the Pareto optimum. □

Proof of Property 4

This is a consequence of Properties 2 and 3. □

Proof of Property 5

If the inequality $F_L \geqslant F_H^M$ holds, then the equality $t_H = t_L$ must be true. This is due to Property 2 and also because the portion of the indifference curve $V(F, t; p_H) = \bar{V}_H$ becomes horizontal as $F \geqslant F_H^M$. To satisfy the IC constraint for type-p_L taxpayers, we need to maintain the inequality $F_L \leqslant F_H$. In fact, as long as $F_L \leqslant F_H$ holds, F_H can be arbitrarily chosen at any point in the horizontal portion of the indifference curve $V(F, t; p_H) = \bar{V}_H$. This result is due to Property 3 and because raising F_H beyond F_H^M will not increase the revenue collected from type-p_H taxpayers. Thus, separating equilibria (i.e. $t_H = t_L$ but $F_L < F_H$) are possible in such a situation. However, all such separating equilibria with $F_L < F_H$ can equivalently be represented by the pooling equilibrium with $F_L = F_H$. As a result, we can ignore these separating cases without loss of generality. We conclude that the Pareto-efficient fine/tax rate structure will be pooling with $(F_H, t_H) = (F_L, t_L) \quad \text{if } F_L \geqslant F_H^M$. □

ACKNOWLEDGMENTS

We are grateful to two anonymous referees for their very valuable comments and suggestions. C. C. Yang gratefully acknowledges the financial support provided by the NSC, Taiwan, under the grant NSC91–2415-H-001–029.

NOTES

1. See Cowell (1990), Andreoni *et al.* (1998) and Slemrod and Yitzhaki (2002) for literature surveys on tax evasion.
2. See Laffont and Martimort (2002, pp. 123–4).
3. Several papers have explored the implications of different probabilities of evasion detection among taxpayers. These include Sandmo (1981); Watson (1985); Kesselman (1989); and Macho-Stadler and Perez-Castrillo (1997). The foci of these papers are quite different from ours; in particular, none of them explores efficient fine/tax rate structures as we do in this paper.
4. According to Mas-Colell *et al.* (1995, pp. 445–6), 'An allocation that cannot be Pareto improved by an authority who is unable to observe agents' private information is known as a constrained (or second-best) Pareto optimum.'
5. Schneider and Enste (2000) provide a recent review of underground economies around the world.
6. Following Schroyen (1997), the amount of income evasion could be interpreted as the fraction of the total income earned in the underground economy. Alternatively, the tax agency may not know *a priori* how difficult or easy it is to provide evidence of evasion for a random chosen taxpayer.
7. See Slemrod and Bakija (1996, chapter 5) and Andreoni *et al.* (1998).
8. This setup is important as far as this paper is concerned. It is arguable that the tax code itself may be imprecise and the tax auditors may not be uniform, so that the so-called 'true

income' may never be known; see Andreoni *et al.* (1998) and Ueng and Yang (2001) for more on this topic.

9. See also Yitzhaki (1987).
10. In the case of $U(c) = \ln c$, we have $v(F, t; p, Y) = V(F, t; p) + g(Y)$.
11. Following a similar argument as in Reinganum and Wilde (1985, fn. 1), one can show that the revelation principle applies in our context. The basic argument runs as follows. If the fine/tax rate policy $(F(\cdot), t(\cdot))$ results in the optimal reporting function $\mathrm{P}(p)$, then the policy given by the composite function $(F(\mathrm{P}(\cdot)), t(\mathrm{P}(\cdot)))$ has an optimal reporting function which induces truthful reporting in detection probability p. This result enables us to confine our attention to the direct revelation mechanism.
12. $F_H = F_H^M$ and $F_L = F_L^M$ will also do the job here.
13. At the Pareto-efficient fine/tax rate structure (see Figure 2 for the separating case and Figure 3 for the pooling case), the slope of the indifference curve for type-p_H taxpayers is horizontal ($x = 1$), while that for type-p_L taxpayers is not ($x < 1$). The case where $F_L \geqslant F_L^M$ corresponds to assigning $\bar{V}_H = \bar{V}_H^M$ in programme (α); this case is atypical and will be discussed later.
14. We will provide an intuition behind Corollary 1 when we discuss the maximum-fine solution later.
15. Commitment is important in our framework. The government will know who the evaders are (i.e. those taxpayers who have a low probability of evasion detection) whenever the separating equilibrium occurs. However, as is typical in a static principal–agent model, the government (principal) is assumed to commit to its announced fine/tax rate structure without utilizing the revealed information of taxpayers (agents) further in the mechanism design.
16. As explained in note 11, the direct revelation mechanism is applicable in our framework. Thus, after the imposition of the IC constraint, taxpayers honestly report their detection probabilities. It does not seem very realistic, however, that the object of reporting by taxpayers is their 'detection probability'. Fortunately, it is not difficult to show that this 'direct revelation mechanism' can be implemented by means of a nonlinear fine/tax rate schedule. The implementation idea is basically the same as that presented by Stiglitz (1987, figure 2.7) for the case of income taxation without evasion. In the Stiglitz case, the object of reporting by taxpayers is their 'ability to earn' if the direct revelation mechanism is involved.
17. Reducing tax rates is ruled out here, since it will violate the IC constraint for type-p_H taxpayers. Also note that reducing fines intended for type-p_L taxpayers will not change the welfare of type-p_H taxpayers as long as $F_H = F_L \geqslant F_H^M$ at the pooling.
18. The best outcome that type-p_H taxpayers can hope for is that $V(F_L, t_L; p_L) = V(F_H, t_H; p_H)$. This is intuitive. A lower p taxpayer can mimic a higher p taxpayer simply by revealing more information to the government. This is not true conversely.
19. It is worth noting that our result is similar to the 'abnormal' case in Stiglitz (1982).
20. The seminal work on the issue is Becker (1968), who suggests that the optimal policy to deter crime is to set fines at their maximum so as to economize on the enforcement cost.
21. Suppose that p_H- and p_L-type taxpayers all have the same level of income. Then it is likely that the government will choose a fine/tax rate structure so that the welfare gap between $V(F_L, t_L; p_L)$ and $V(F_H, t_H; p_H)$ is as small as possible. This may lead to the maximum-fine policy with $F_H = F_L = F_L^M$. Now consider a more plausible case, in which p_H- and p_L-type taxpayers have different levels of income (say, Y_L and Y_H). The smallest gap between $V(F_L, t_L; p_L)$ and $V(F_H, t_H; p_H)$ might no longer be a desirable goal in such a situation. For example, if the p_H-taxpayers (compliers) have a much higher income on average than the p_L-type taxpayers (evaders), it might be better to widen rather than narrow the gap between $V(F_L, t_L; p_L)$ and $V(F_H, t_H; p_H)$ so that the gap between $v(F_L, t_L; p_L, Y_L) = V(F_L, t_L; p_L) \cdot g(Y_L)$ and $v(F_H, t_H; p_H, Y_H) = V(F_H, t_H; p_H) \cdot g(Y_H)$ can become smaller. This may then lead to a policy associated with $\bar{V}_H < \bar{V}_H^M$ and hence to a fine less than the maximum. Of course, a tradeoff could be involved if the government took into account other criteria (say, that tax evasion will not be tolerated in any circumstances). Note that a less-than-maximum-fine policy is consistent with the observation in the real world that governments typically do not set fines high enough to deter all evasions (see Andreoni *et al.* 1998, pp. 823–4).
22. For the arguments of NNWE, see Stiglitz (1987).
23. There does exist some evidence in support of the CRRA assumption (see the references cited in Yaniv 1994).

REFERENCES

ALLINGHAM, M. G. and SANDMO, A. (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics*, **1**, 323–38.

ANDREONI, J., ERARD, B. and FEINSTEIN, J. (1998). Tax compliance. *Journal of Economic Literature*, **36**, 818–60.
ARROW, K. J. (1970). *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland.
BALDRY, J. C. (1984). The enforcement of income tax laws: efficiency implications. *Economic Record*, **60**, 156–59.
BECKER, G. (1968). Crime and punishment: an economic approach. *Journal of Political Economy*, **76**, 169–217.
BORDER, K. C. and SOBEL, J. (1987). Samurai accountant: a theory of auditing and plunder. *Review of Economic Studies*, **54**, 525–40.
CHANDER, P. and WILDE, L. L. (1998). A general characterization of optimal income tax enforcement. *Review of Economic Studies*, **65**, 165–83.
COWELL, F. A. (1985). The economic analysis of tax evasion. *Bulletin of Economic Research*, **37**, 163–93.
——— (1990). *Cheating the Government: The Economics of Evasion*. Cambridge, Mass.: MIT Press.
CREMER, H. and GAHVARI, F. (1995). Tax evasion and the general income tax. *Journal of Public Economics*, **60**, 235–49.
KESSELMAN, J. R. (1989). Income tax evasion: an intersectoral analysis. *Journal of Public Economics*, **38**, 137–82.
LAFFONT, J. J. and MARTIMORT, D. (2002). *The Theory of Incentives: The Principal–Agent Model*. Princeton: Princeton University Press.
MACHO-STADLER, I. and PEREZ-CASTRILLO, D. J. (1997). Optimal auditing with heterogeneous income sources. *International Economic Review*, **38**, 951–68.
MAS-COLELL, A., WHINSTON, M. D. and GREEN, J. R. (1995). *Microeconomic Theory*. Oxford: Oxford University Press.
MIRRLEES, J. A. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies*, **38**, 175–208.
——— (1975). Optimal commodity taxation in a two-class economy. *Journal of Public Economics*, **4**, 27–33.
MOOKHERJEE, D. and PNG, I. P. L. (1989). Optimal auditing, insurance, and redistribution. *Quarterly Journal of Economics*, **104**, 399–415.
REINGANUM, J. F. and WILDE, L. L. (1985). Income tax compliance in a principal–agent framework. *Journal of Public Economics*, **26**, 1–18.
SANDMO, A. (1981). Income tax evasion, labour supply, and the equity–efficiency tradeoff. *Journal of Public Economics*, **16**, 265–88.
SCHNEIDER, F. and ENSTE, D. (2000). Shadow economies: size, causes, and consequences. *Journal of Economic Literature*, **38**, 77–114.
SCHROYEN, F. (1997). Pareto efficient income taxation under costly monitoring. *Journal of Public Economics*, **65**, 343–66.
SLEMROD, J. and BAKIJA, J. (1996). *Taxing Ourselves*. Cambridge, Mass.: MIT Press.
SLEMROD, J. and YITZHAKI, S. (2002). Tax avoidance, evasion, and administration. In A. J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Vol. 3. Amsterdam: North-Holland.
STIGLITZ, J. (1982). Utilitarianism and horizontal equity: the case for random taxation. *Journal of Public Economics*, **18**, 1–33.
——— (1987). Pareto efficient and optimal taxation and the New Welfare Economics. In A. J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, Vol. 2. Amsterdam: North-Holland.
UENG, K. L. and YANG, C. C. (2001). Plea bargaining with the IRS: extension and further results. *Journal of Public Economics*, **81**, 83–98.
WATSON, H. (1985). Tax evasion and labour markets. *Journal of Public Economics*, **27**, 231–46.
WEISS, L. (1976). The desirability of cheating incentives and randomness in the optimal income tax. *Journal of Political Economy*, **84**, 1343–52.
YANIV, G. (1994). Tax evasion and the income tax rate: a theoretical re-examination. *Public Finance*, **49**, 107–12.
YITZHAKI, S. (1974). A note on income tax evasion: a theoretical analysis. *Journal of Public Economics*, **3**, 201–2.
——— (1987). On the excess burden of tax evasion. *Public Finance Quarterly*, **15**, 123–37.