Plea bargaining with the IRS: extensions and further results

K.L. Glen Ueng[a], C.C. Yang[b,*]

[a]*Department of Economics, National Chi Nan University, Puli, Nantou Hsien 545, Taiwan*
[b]*Institute of Economics, Academia Sinica, Nankang, Taipei 115, Taiwan*

Received 31 August 1999; accepted 29 February 2000

Abstract

Chu [Journal of Public Economics, 41 (1990) 319–333] proposes a tax scheme called FATOTA, under which taxpayers are given the choice between paying a fixed amount of taxes and thus being exempted from tax audit, or paying only what taxpayers claim they owe and being subject to tax audits for possible evasion. In this paper we extend Chu's finding with regard to the Pareto-improving property of FATOTA to a more general setting, including *nonlinear* tax/penalty schedules, *complicated* audit rules, and *endogenous* labor supply. We also show that any tax system in which the highest-income taxpayer faces audit risk and/or tax distortion can be Pareto-improved upon by the introduction of FATOTA. We relate this result to Seade's [Journal of Public Economics, 7 (1977) 203–235] celebrated finding in the optimal taxation literature that any income tax schedule with a *positive* marginal tax rate at the top of the scale can be replaced by another tax schedule which is strictly Pareto-improving. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Plea bargaining; Tax evasion; Tax reform; Optimal taxation

JEL classification: H21; H26

1. Introduction

Chu (1990) proposes a tax scheme called FATOTA, under which taxpayers are

*Corresponding author. Tel.: +886-2-2782-2791; fax: +886-2-2785-3946.
E-mail address: ccyang@ieas.econ.sinica.edu.tw (C.C. Yang).

given the choice between paying a fixed amount of taxes (FAT); set by the tax authority and thus being exempted from tax audit, or paying only what taxpayers claim they owe and being subject to tax audits (TA); for possible evasion. This scheme is analogous to the plea bargaining system of the US criminal proceedings, as pointed out by Chu. Under the assumptions of risk-averse taxpayers, linear tax and penalty schedules, random audit rules, and exogenous labor supply, Chu applies the Pareto criterion to study a tax reform problem: introducing FATOTA to the present tax system. He shows that a *marginal* introduction of FATOTA will be Pareto-improving, leaving no taxpayers worse off but resulting in a non-trivial increase in government revenue.[1]

In this paper we stick to the assumption that taxpayers are risk averse, but consider a setting more general than Chu's. This includes *nonlinear* tax/penalty schedules, *complicated* audit rules, and *endogenous* labor supply. We prove in our general setting that a *discrete* introduction of FATOTA is still Pareto-improving.

As emphasized above, our model allows for endogenous labor supply. This set-up would be consistent with the literature on optimal income taxation pioneered by Mirrlees (1971) and, more importantly, would enable us to integrate our analysis into the optimal taxation literature.

In the optimal taxation literature, Seade (1977) has a celebrated result concerning the shape of the optimal tax schedules. He shows that any income tax schedule with a *positive* marginal tax rate at the top of the scale can be replaced by another tax schedule which is strictly Pareto-improving. This result is in striking contrast to actual tax schedules and negates in a blunt way the general presumption that marginal tax rates should rise with income.

In this paper we derive a result related to Seade's. It is shown that any tax system in which the highest-income taxpayers face audit risk and/or tax distortion can be Pareto-improved upon by the introduction of FATOTA. We compare this result with Seade's.

The rest of the paper is organized as follows. Section 2 introduces the basic model. Section 3 reports the findings from our analysis. Section 4 extends the basic model to include leisure. Sections 5 and 6 provide discussion and concluding remarks.

2. The basic model

Consider an economy in which there are $n < \infty$ taxpayers. Each taxpayer is characterized by an income level $Y \in (0, \infty)$, which is private information and unknown to the tax authority (without auditing). All taxpayers are assumed to have

[1] For the precise meaning of the 'marginal introduction' of FATOTA, see later or Chu (1990).

identical preferences as in Chu (1990). Our main results remain robust with extension to heterogeneous preferences (see later).

Taxpayers are assumed to be risk averse. As to the risk attitudes of the tax authority toward revenue, we exclude the possibility of risk loving. The tax authority is normally assumed to be risk neutral toward revenue in the literature. However, as argued in Yang (1993), there are situations where assuming a risk-averse tax authority may be more appropriate.

Let the tax system in operation be represented by an income tax schedule $T(.)$ with $0 < T' < 1$ (i.e. positive but less than 100% marginal tax rates), a penalty function $F(.)$ with $F' \geq 0$ (non-decreasing fines), and an audit probability $p > 0$ (i.e. the random audit rule independent of taxpayers' reported income). The penalty function is assumed to depend on the amount of taxes evaded and may be nonlinear. The audit rule is assumed to be random as in Chu (1990), but this assumption can be relaxed (see later). The tax system in question will be called the regular tax system for later reference.

Faced with the regular tax system, a taxpayer (characterized by Y_i) is assumed to report income X_i ($\leq Y_i$) so as to maximize the expected utility:

$$EU(Y_i) \equiv (1 - p) \cdot U[Y_i - T(X_i)] + p \cdot U\{Y_i - T(X_i) - F[T(Y_i) - T(X_i)]\} \qquad (1)$$

where $U(\cdot)$ is a von Neumann-Morgenstern utility function with $U' > 0$ and $U'' < 0$. The set-up of (1) is in the tradition of decision under risk, and has been popular in the tax evasion literature since the seminal work of Allingham and Sandmo (1972) and Yitzhaki (1974). The set-up assumes: (i) true income will be discovered once the tax evader gets audited, and (ii) the caught evader will be fined, and a penalty levied on the amount of evaded tax, as is the case under most tax laws. Apart from allowing for nonlinear tax/penalty schedules, this model is basically the same as in Chu (1990).

Given the regular tax system, a taxpayer will engage in tax evasion whenever evasion is a favorable gamble to the taxpayer. This is because a risk averter takes no part of an unfavorable or barely fair gamble, but always takes some part of a favorable gamble (see Arrow (1970, pp. 99–100)). In what follows we assume that tax evasion is a favorable gamble such that all taxpayers will evade tax. This is a simplified assumption for ease of exposition. The same assumption is also imposed in Chu (1990, Proposition 1). In Section 5 we extend our model to account for the case where some taxpayers do not evade tax.

As in Chu (1990), our focus is on the introduction of FATOTA to complement the regular tax system.

3. The analysis

Faced with the regular tax system, all taxpayers choose to evade by assumption. However, a taxpayer who evades tax also exposes himself to audit risk. The maximal amount of money, β_i, that taxpayer i would be ready to hand over to the tax authority for exemption from the regular tax system is defined implicitly by

$$U(Y_i - \beta_i) = EU(Y_i). \tag{2}$$

All the β_is must be positive in view of the risk-averse assumption that $U'' < 0$. Clearly, the value of β_i depends upon the parameters of the regular tax system as well as upon income Y_i.

Suppose that the government offers a FATOTA program to the taxpayers. The existence of this program means that those choosing to pay FAT will be exempted from the regular tax system, while those choosing to face TA will stay with the regular tax system. Given the either-or choice between FAT and TA, Chu's Proposition 1 states that there exists a threshold value of income such that a taxpayer will choose FAT instead of TA if his income is above the threshold. Chu relies upon this result in his proof of the Pareto-improving property of FATOTA. Using (2) and the envelope theorem, we have:

$$\frac{\partial \beta_i}{\partial Y_i} = 1 - \frac{(1-p)\cdot U'[Y_i - T(X_i)] + p\cdot U'\{Y_i - T(X_i) - F[T(Y_i) - T(X_i)]\}\cdot(1 - T'F')}{U'(Y_i - \beta_i)} \tag{3}$$

the sign of which is ambiguous.[2] A taxpayer's 'willingness to pay' for the exemption from the regular tax system or TA is not necessarily monotonic in income as the sign of (3) is ambiguous. As a result, Chu's Proposition 1 may not hold under our general setting. This in turn implies that we need to seek a method different from Chu's to prove the Pareto-improving property of FATOTA. Our proof is as follows.

The expected tax revenue (including fines) collected from taxpayer i when the regular tax system is enforced equals

$$ER_i = (1-p)\cdot T(X_i) + p\cdot\{T(X_i) + F[T(Y_i) - T(X_i)]\} \tag{4}$$

and so

$$Y_i - ER_i = (1-p)\cdot[Y_i - T(X_i)] + p\cdot\{Y_i - T(X_i) - F[T(Y_i) - T(X_i)]\}. \tag{5}$$

Applying Jensen's inequality with $U'' < 0$ to (5) yields

$$U(Y_i - ER_i) > (1-p)\cdot U[Y_i - T(X_i)] + p \cdot U\{Y_i - T(X_i) - F[T(Y_i) - T(X_i)]\} = EU(Y_i). \tag{6}$$

From (2) and (6), we observe that

$$\beta_i > ER_i. \tag{7}$$

---

[2]If both the tax and the penalty schedule are linear, $T'F' > 1$ as shown in the Appendix of Chu (1990). Then the sign of (3) will be positive (note that $U'(Y_i - \beta_i) > U'[Y_i - T(X_i)]$ since $(Y_i - \beta_i) < [Y_i - T(X_i)]$).

This demonstrates that each taxpayer would be ready to pay more than the expected 'tax' levied on him by the regular tax system.

From (4),

$$\frac{\partial ER_i}{\partial Y_i} = (1-p)\cdot T' \frac{\partial X_i}{\partial Y_i} + p\cdot\left[T' \frac{\partial X_i}{\partial Y_i} + F'T'\left(1-\frac{\partial X_i}{\partial Y_i}\right)\right]. \tag{8}$$

It can be shown that reported incomes will rise with true incomes but the marginal increase will not exceed 1 (i.e. $0 < (\partial X_i / \partial Y_i) < 1$). The sign of (8) is thus positive: the higher the true income, the higher the expected revenue raised.

Following Chu (1990), let FATOTA(x) represent the FATOTA system imposing a FAT x. From the definition of β_i, it is clear that if the taxpayers are given the choice between paying FAT or facing TA under FATOTA(x), all taxpayers having $\beta_i > x$ would choose to pay the FAT x.

Now, consider the taxpayer who has the maximal income $Y_{\max} = \max\{Y_1, Y_2, \ldots, Y_n\}$, and let $\beta(Y_{\max})$ and $ER(Y_{\max})$ denote his 'willingness to pay' and 'expected revenue contribution', respectively. According to the sign of (8), $ER(Y_{\max}) = \max\{ER_1, ER_2, \ldots, ER_n\}$. From (7) we also have $\beta(Y_{\max}) > ER(Y_{\max})$. Suppose that the taxpayers are given the choice between paying FAT or facing TA under FATOTA(x), where $ER(Y_{\max}) < x < \beta(Y_{\max})$. Then the taxpayer characterized by $Y_{\max}$ will surely choose to pay the FAT x. Other taxpayers having $\beta_i > x$ will also choose to pay the FAT x. It is obvious that all of the taxpayers choosing FAT will be better off, while the rest choosing TA will not be worse off since they simply stay with the regular tax system. As to the change in government revenue, we differentiate the taxpayers into two groups: those choosing FAT and those choosing TA. The tax authority will raise the same expected amount of revenue from the second group. But it will raise more than before from the first group since all the taxpayers choosing FAT pay a FAT x with $x > ER(Y_{\max}) \geq ER_i$ for all i. This completes the proof of the following result.

Proposition 1. *Let the regular tax system be in effect at the status quo. A discrete introduction of FATOTA (at $ER(Y_{\max}) < x < \beta(Y_{\max})$) will be Pareto-improving.*

The term 'discrete introduction' in the above proposition is used to contrast with the 'marginal introduction' of FATOTA in Chu (1990). Consider FATOTA($\beta_{\max}$), where $\beta_{\max} = \max\{\beta_1, \beta_2, \ldots, \beta_n\}$. Given the choice between FAT and TA under FATOTA($\beta_{\max}$), all taxpayers other than the one having $\beta_{\max}$ will choose TA. According to the definition of $\beta_{\max}$, the taxpayer having $\beta_{\max}$ is actually indifferent between FAT and TA under FATOTA($\beta_{\max}$). However, let the taxpayer choose to pay the FAT $\beta_{\max}$. This choice follows the convention in the principal-agent model that agents (taxpayers) are assumed to choose the course of action the principal (tax authority) desires if they are indifferent between courses of action. Then, because of $\beta_{\max} \geq \beta(Y_{\max}) > ER(Y_{\max}) = \max\{ER_1, ER_2, \ldots,$

$ER_n\}$, we immediately see that the introduction of FATOTA($\beta_{\max}$) causes no taxpayers to become worse off but results in a non-trivial increase in government revenue. This is basically the welfare outcome of the 'marginal' introduction of FATOTA analyzed in Chu (1990) Proposition 2. Proposition 1's 'discrete' introduction of FATOTA is a slight innovation relative to Chu's Proposition 2 in the sense that our Pareto-improving introduction of FATOTA(x) need not be confined to the case where $x = \beta_{\max}$. Note finally that neither in the 'discrete' nor 'marginal' introduction do we rely upon Chu's Proposition 1 in our proof.[3]

Remark 1. *Under* (1), *the audit rule is random and taxpayer preferences are identical. It should be clear however that more complicated audit rules and/or heterogeneous preferences could be allowed in our framework. The validity of Proposition* 1 *relies upon uncertainty induced by the regular tax system. It does not rely upon the random audit rule, nor does it rely upon identical preferences.*

Remark 2. *One can rewrite* (6) *as*

$$U(Y_i - ER_i - \theta_i) = EU(Y_i), \tag{9}$$

where $\theta_i > 0$ *is the risk premium that taxpayer i would be ready to pay in order to eliminate the exposure to audit risk. As interpreted by Chu* (1990), *the marginal introduction of FATOTA is in effect 'confiscating' risk premium from the taxpayer having the maximal 'willingness to pay'. Yitzhaki* (1987) *calls the risk premium defined in* (9) *the 'excess burden of tax evasion'. It is an excess burden because* θ_i *represents a deadweight loss beyond what would be imposed if* ER_i *were somehow collected by a lump-sum tax. The existence of this excess burden under the regular tax system is the key reason why the introduction of FATOTA can Pareto-improve upon the regular tax system. Our proof of Proposition* 1 *exploits the existence of the excess burden of tax evasion directly.*

4. An extended model with leisure

Since income is a parameter in the basic model, taxation under the regular tax system does not allow for any distortion caused by altering labor supply. The labor supply question, however, lies at the heart of the optimal income taxation literature. In this section we extend the basic model to include leisure so as to integrate our analysis into the optimal taxation literature.

---

[3] Chu also derives the optimal FATOTA system through numerical examples. The optimal criterion used is with respect to a social welfare function. We do not address the optimality here, for no more insight into the issue than Chu's is obtained. See also Chu (1990) for the practical problems associated with applying the FATOTA system to corporations.

4.1. The extended model

The features of the extended model remain the same as those in the basic model, except for the following changes. There are two commodities: a consumption good c and labor l. All taxpayers have a common utility function denoted by $U(c, l)$ with $U_c > 0$, $U_l < 0$, and $U_{cc} < 0$. Each taxpayer is characterized by ability or wage rate $w \in (0, \infty)$, which is private information and known only to the taxpayer himself. Taxpayer i supplies labor l_i and hence earns pre-tax income $Y_i = w_i l_i$.

4.2. The analysis

Faced with the regular tax system, a taxpayer (characterized by w_i) is assumed to report income X_i ($\leq Y_i$) so as to maximize the expected utility:

$$EU(w_i) \equiv (1 - p(.)) \cdot U[Y_i - T(X_i), l_i] + p(.) \cdot U\{Y_i - T(X_i) - F[T(Y_i) - T(X_i)], l_i\} \quad (10)$$

where $p(.)$ denotes some audit rule. The maximal amount of money, β_i, that taxpayer i would be ready to hand over to the tax authority for exemption from the regular tax system is now defined implicitly by

$$U[w_i l_i^*(\beta_i) - \beta_i, l_i^*(\beta_i)] = EU(w_i), \quad (11)$$

where the left-hand side is an indirect utility function. After making the lump-sum payment β_i so as to be exempted from the regular tax system, taxpayer i would choose the labor, l_i^*, that maximizes utility. This is the reason why l_i^* is specified as a function of β_i in Eq. (11).

With all the β_is at hand, we can proceed with the analysis in the same way as in the last section. There is an additional result, however. Let us define β_i^0 implicitly in terms of the following equality

$$U(w_i l_i^0 - \beta_i^0, l_i^0) = EU(w_i), \quad (12)$$

where l_i^0 is the labor chosen by taxpayer i under the regular tax system. Using (5) and applying Jensen's inequality with $U_{cc} < 0$ yields

$$U(Y_i^0 - ER_i, l_i^0) > (1 - p(.)) \cdot U[Y_i^0 - T(X_i), l_i^0] + p(.) \cdot U\{Y_i^0 - T(X_i) - F[T(Y_i^0) - T(X_i)], l_i^0\} = EU(w_i) \quad (13)$$

where $Y_i^0 = w_i l_i^0$. From (12) and (13),

$$\beta_i^0 > ER_i. \quad (14)$$

Since l_i^* is the labor that maximizes utility after the payment β_i is made, we obtain[4]

$$\beta_i > \beta_i^0 > ER_i \quad \text{if} \quad l_i^* \neq l_i^0. \tag{15}$$

Intuitively, l_i^0 in (12) is a constrained labor supply while l_i^* in (11) is not. The utility derived from l_i^0 is therefore expected to be lower than that derived from l_i^*. To make the left-hand side of (12) equal to that of (11), it is logical to see that $\beta_i > \beta_i^0$. Alternatively, β_i acts like a lump-sum tax while β_i^0 is obtained under the income tax $T(.)$ with tax distortion as a rule. The tax distortion is reflected by changes in the choice of labor supply from l_i^* to l_i^0. The gap between β_i and β_i^0 measures the usual excess burden of tax distortion, if there is any.

If we define the risk premium as in (9), then from (12) and (13)

$$U(w_i l_i^0 - ER_i - \theta_i, l_i^0) = U(w_i l_i^0 - \beta_i^0, l_i^0) = EU(w_i). \tag{16}$$

This result leads to

$$\beta_i^0 = ER_i + \theta_i \tag{17}$$

and hence

$$\beta_i = ER_i + \theta_i + (\beta_i - \beta_i^0). \tag{18}$$

Taxpayer i's willingness to pay for the exemption from the regular tax system exceeds ER_i (the expected revenue) by the amount θ_i (the excess burden of tax evasion) and $\beta_i - \beta_i^0$ (the excess burden of tax distortion).

To sum up, in addition to the excess burden of tax evasion, there exists the excess burden of tax distortion under the regular tax system if labor supply decisions are endogenous. The existence of these two types of excess burden under the regular tax system makes room for the welfare-improving possibility.

Consider FATOTA(x), where $ER(Y_{\max}) < x < \beta(Y_{\max})$ and $Y_{\max} = \max\{Y_1, Y_2, \ldots, Y_n\}$ in the extended model. Then the same argument as before applies and so Proposition 1 holds even after the incorporation of endogenous labor supply.

Assume that the taxpayer having $\beta_{\max}$ chooses to pay the FAT under FATOTA($\beta_{\max}$) (i.e. the marginal introduction of FATOTA). Then, in the extended model here, the government is not only 'confiscating' the taxpayer's risk premium but also his benefit resulting from the elimination of tax distortion.

4.3. On optimal taxation of the highest income

In the optimal taxation literature, Seade (1977) has a celebrated result concerning the shape of the optimal tax schedules. He shows that any income tax

[4] See Ueng and Yang (2000) for the details.

schedule with a *positive* marginal tax rate at the top of the scale can be replaced by another tax schedule which is strictly Pareto-improving. This result is derived without taking into account the possibility of tax evasion and of introducing FATOTA to the tax system. In what follows we reconsider Seade's result in our setting.

Consider the taxpayer having the maximal pre-tax income $Y_{max} = \max\{Y_1, Y_2, \ldots, Y_n\}$ under the regular tax system. Exploiting the property that $\beta(Y_{max}) > ER(Y_{max}) \geq ER_i$ for all i, the government offers the taxpayers FATOTA(x) with $ER(Y_{max}) < x < \beta(Y_{max})$. Then the same argument as that leading to Proposition 1 applies and so the introduction of FATOTA(x) will be Pareto-improving. Note that the inequality $\beta(Y_{max}) > ER(Y_{max})$ holds as long as there exists an excess burden of tax evasion or an excess burden of tax distortion imposed upon the highest income taxpayers. We state formally:

Proposition 2. *Any tax system in which the highest income taxpayers face audit risk and/or tax distortion can be Pareto-improved upon by the introduction of FATOTA.*

The key to Proposition 2, of course, lies in the fact that the introduction of FATOTA can eliminate the excess burden of both tax evasion and tax distortion that has been imposed. The elimination of the excess burden of tax evasion is easy to understand since there will be no TA after the payment of FAT. As to the elimination of the excess burden of tax distortion, this is also easy to understand since a FAT acts like a lump-sum tax.

The statement of Proposition 2 is restricted to the highest income taxpayers. This is because other taxpayers' paying FAT may result in a loss in government revenue and so the outcome may not be Pareto-improving. Consider FATOTA(x) with $ER(Y_i) < x < \beta(Y_i)$ for some i and $Y_i \neq Y_{max}$. Since the inequality $x < ER(Y_{max})$ may hold, we cannot exclude the possibility that the introduction of FATOTA(x) may reduce expected government revenue.

Note that a FAT implies a zero marginal tax rate, but the converse need not be true. This observation combined with Proposition 2 suggests that, even if the highest income taxpayers face the zero marginal tax rate, the tax system can be Pareto-improved further by the introduction of FATOTA as long as audit risk and/or tax distortion are imposed on the highest income taxpayers. In this sense our tax reform is stronger than Seade's.

In another sense, however, our tax reform is weaker than Seade's. Although both tax reforms are Pareto-improving, Seade's is 'non-local' while ours is 'local'. Seade shows that it is possible to construct another tax schedule to replace the one with a positive marginal tax rate at the top such that all or most taxpayers will be better off. Because the sign of (3) is ambiguous, some taxpayers may benefit from the introduction of FATOTA even if they are not characterized by the highest income. Overall, however, our Pareto-improving FATOTA only guarantees to increase the utilities of the very rich or their tax payments, leaving most other taxpayers stuck with the regular tax system.

To have a FATOTA Pareto improvement such that the benefit is shared by the taxpayers at large, one possibility is to let the tax authority offer a FAT to a group of taxpayers as a whole rather than individually. As long as the taxpayers in question can raise enough money *as a group* to meet the FAT set by the tax authority, each and every taxpayer in the group will be exempted from the regular tax system and hence exempted from the excess burden of tax evasion and/or tax distortion imposed. This possibility has been explored in Ueng and Yang (2000).

5. Discussion

Our previous analysis takes as its point of departure the premise that cheating incentives exist such that all taxpayers evade tax. This set-up is convenient but unrealistic. For example, according to Roth et al. (1989), 58% of all tax returns had no understatement of taxable income in an IRS-sponsored study based on the 1979 Taxpayer Compliance Measurement Program. In this section we examine the robustness of our previous results under the extension that some people do not evade tax.

A possible way to extend our model is to follow the idea of Watson (1985) and Macho-Stadler and Perez-Castrillo (1997). They basically assume that amounts of information available to the tax authority are different between taxpayers and, as a result, different taxpayers in a sense have different abilities to hide their true incomes (for example, it is easier for the evasion of a wage/salary earner to be detected by the tax authority than that of a self-employed). Let e denote the tax authority's audit efforts and k denote the taxpayer type with regard to the ability to hide the true income. Then the audit probability p can be simply redefined as $p = e \cdot k$. For simplicity, consider the case of a constant fine F. If a taxpayer under-reports a dollar of tax, he will gain one dollar if evasion succeeds, but with probability p will lose F if evasion fails. Therefore, the expected value of each dollar of tax cheating is $1-pF$. Let $e\hat{k}F = 1$ so that the taxpayer with $k = \hat{k}$ faces a fair gamble of tax evasion. We then obtain the result that the taxpayers with $k < \hat{k}$ (facing a favorable gamble) will evade tax while those with $k \geq \hat{k}$ (facing a fair or unfavorable gamble) will not.[5]

The extension above, however, critically relies upon the constraint of fines. It is clear that if fines are high enough, then no taxpayer will evade according to our

[5] Alternatively, one may assume that different taxpayers incur different 'psychic costs' as in Gordon (1989), or have different degrees of 'conformities to social customs' as in Myles and Naylor (1996). This extension will also generate the result that some people evade tax while others do not. We shall not present this line of extension in order to save space. Our arguments in the text basically apply to this extension as well.

analysis. Raising fines to the possible maximum appears to be an easy way to eliminate tax evasion since it involves little cost in general. Indeed, a classical argument by Becker (1968) is that, to economize on enforcement costs, fines should be set as high as possible, with the corresponding probability of apprehension being as low as possible. Governments generally do not set sanctions very high in reality, however.[6] Andreoni et al. (1998) document that the IRS typically applies civil penalties at a rate of 20% of the portion of the underpayment of tax; in cases of fraud, a civil penalty may be applied at the rate of 75%.[7] These penalties obviously are not at their maximum.

The focus of this paper is on the issue of tax reform (the Pareto-improving property of introducing FATOTA to the existing tax system), as distinct from de novo tax design (writing the tax laws on 'a clean sheet of paper' as it was put by Feldstein (1976)). In view of our *reform* rather than *design* focus, we shall not explore the optimality issue as to why fines in reality are constrained.[8] Instead, we shall take the constrained fine in the regular tax system as a given fact and allow for evasion by some, but not all, of the taxpayers as in our extension.

Given the constrained fine, we now discuss whether the highest income taxpayers tend to evade and whether the highest income taxpayers benefit from the introduction of FATOTA. This discussion is important since it is the highest income taxpayers who opt for FAT in our model.[9]

Note first that, given the audit effort e and the fine F, whether or not a taxpayer will evade tax is determined solely by his ability to evade k (i.e. if the inequality $ekF < 1$ holds). We do not know if the ability to evade and the level of income are positively or negatively correlated. Nevertheless, most empirical studies support a positive relationship between income and evasion.[10] On the basis of this piece of evidence, we may not be able to conclude that the highest income taxpayers evade more than others. But the possibility that the highest income taxpayers do evade cannot be ruled out, at least a priori.

---

[6]Since the seminal paper of Becker (1968), a large volume of literature has focused on offering explanations for why the use of maximal sanctions may be inefficient. See Bebchuk and Kaplow (1993, Footnote 1) for a brief summary of these explanations. More recent contributions to the issue include Andreoni (1991), Franzoni (1999) and Friedman (1999).

[7]In cases of felony, tax evasion is punishable by a fine of not more than $100,000, imprisonment for not more than 5 years, or both. Severe penalties are quite infrequently imposed, however (see Andreoni et al. (1998)).

[8]Kolm (1973), Sandmo (1981), Andreoni (1992) and Pestieau et al. (1994) have all shown in the context of tax evasion that maximal fines may not be optimal.

[9]Our proof of Propositions 1 and 2 uses the monotonic relationship between income and expected tax revenue as described by Eq. (8). It is obvious however that, as long as the highest income taxpayers contribute the highest expected tax revenue, Propositions 1 and 2 will remain true. We implicitly make this assumption here. When some taxpayers do not evade, the monotonic relationship between income and expected tax revenue may no longer hold, but that the highest income taxpayers contribute the highest expected tax revenue seems still plausible.

[10]See Andreoni et al. (1998) for a summary of empirical evidence.

One may argue that the highest income taxpayers often seek the help of professional tax practitioners and rarely file returns by themselves. This argument could dilute the view that the highest income taxpayers tend to evade. However, the empirical study by Erard (1993) finds that the use of attorneys or certified public accountants to prepare returns is significantly associated with increased rather than decreased noncompliance. Klepper et al. (1991) corroborate this finding. They show that although tax practitioners appear to contribute to compliance by reducing the incidence of error, they may actually promote greater noncompliance through reducing the perceived chances of audit and penalty, lowering the psychic and monetary costs associated with audits, or exploiting aggressively ambiguous features of the tax code.

It is assumed so far that true income will be discovered once the tax authority carries out the audit. The validity of this assumption may be questionable, however. The tax code itself may be imprecise, or the quality of tax auditors may not be uniform. Slemrod and Bakaija (1996, Chapter 5) report that *Money* magazine invented a tax situation of moderate, but not exceptional, complexity and asked tax professionals to calculate tax liability. In 1993 they got back 41 different answers from the 41 professionals and the answers ranged from $31,864 to $74,450 (the correct answer was $35,643). The same set of questions was put to several taxpayer service representatives of the IRS. In 1992, 86% of the questions were answered correctly and in 1988 the correct figure was only 55%. That the answers are far short of ideal comes as no surprise. According to Hall (1995), the total number of words in the two documents, *Internal Revenue Code* and *Federal Tax Regulations*, could be over 5.5 million. These pieces of casual evidence to some extent suggest that the true application of tax laws may be impossible to resolve prior to an audit, the idea of 'honest reporting' may have no meaning, and a taxpayer may never be free from audit risk. Several authors, including Alm (1988), Scotchmer (1989) and Scotchmer and Slemrod (1989), have provided formal analyses on this kind of taxpayer uncertainty.[11] As far as this paper is concerned, we would like to emphasize that the uncertainty faced by the risk-averse taxpayers may not be at their discretion. The uncertainty could simply result from possible errors of observation in the assessment of income by the tax authority, or any other risk due to the practice of auditing.

The ultimate reason for taxpayer uncertainty is the complexity of tax laws. The complexity may have its own purpose, but it is costly for taxpayer compliance. Slemrod and Sorum (1984) and Blumenthal and Slemrod (1992) estimate that the total cost for complying with the US individual income tax is between 5% and 7% of revenue raised. These costs include the monetary outlays for professional

[11] This idea of taxpayer uncertainty can be readily incorporated into our model, for example, by viewing incomes or fines as random variables before a tax audit. Our main results carry through in this setting; details are available from the authors upon request.

guidance and the time spent by taxpayers themselves. Since the highest income taxpayers as a rule tend to have more complex returns, it is arguable that the highest income taxpayers are likely to be subject to serious taxpayer uncertainty and/or paying above-average compliance costs even in the absence of tax evasion.

One cannot rule out the possibility a priori that the highest income taxpayers do evade tax under the regular tax system. Even with 'honest reporting', the uncertainty or randomness associated with the regular tax system may be rather high for the highest income taxpayers. The compliance costs borne by the highest income taxpayers are likely to be substantial under the regular tax system. All of these audit risks and/or compliance costs incurred by the highest income taxpayers can be lessened or even eliminated through the introduction of FATOTA.

6. Concluding remarks

As noted by Chu (1990, Footnote 6), the FATOTA system bears some resemblance to the so-called 'cutoff rule' in the design of audit policy. This rule was first introduced into the tax evasion literature by Reinganum and Wilde (1985). In contrast to the random rule, which does not distinguish between taxpayers on the basis of their reported income, the cutoff rule makes use of information contained in the reported income in a simple way. More precisely, it specifies a threshold income level and audits any report below the threshold with some probability, but leaves all reports above the threshold unaudited.

Faced with a cutoff rule that is announced and committed by the tax authority, it is plausible to assume that taxpayers with incomes above the threshold will only report the threshold income. Then a tax system incorporating the cutoff rule is not different from the FATOTA system proposed by Chu. Taxpayers still face the choice between paying a FAT corresponding to the threshold income and thus being exempted from tax audit, or paying only what they claim they owe and being subject to TA for possible evasion. With such an interpretation, our study of FATOTA is essentially a study of the cutoff rule.[12]

Proposition 2 implies that, at the Pareto optimum, the highest income taxpayers must be free from any audit risk. This result is reminiscent of a similar recent finding by Cremer and Gahvari (1995) and Schroyen (1997). Under the setting that there are only two types of taxpayers (high- and low-ability), both papers derive the result that high-ability taxpayers should never be audited when tax evasion is incorporated into an optimum general income tax problem. Suppose that pre-tax income Y_i is increasing in ability w_i (this property may not hold under our

[12] Sanchez and Sobel (1993) provide a detailed discussion of the cutoff rule. The discussion, however, proceeds under the assumption that taxpayers are risk neutral.

setting). Then it should be clear that the audit-free result carries over to the highest-ability taxpayers in fairly general situations. These situations include: (i) there are more than two types of taxpayers, (ii) tax schedules need not be optimal, (iii) incentive constraints typically imposed in the principal-agent model can be ignored, (iv) consumption and leisure need not be assumed to be separable as in Cremer and Gahvari (1995) or Edgeworth complementarity as in Schroyen (1997), and (v) audit risk can stem from the tax authority's errors in assessing income in the absence of tax evasion. Note also that, as long as tax distortion exists, Proposition 2 holds even if there is no audit risk.

The key idea underlying FATOTA is that its introduction offers a way of limiting audit risk and/or tax distortion that taxpayers face under the regular tax system. The magnitude of benefits in introducing FATOTA will thus depend upon how serious audit uncertainty and/or tax distortion are in different situations.

We conclude our paper with two comments. Firstly, even though we have described our economy as if it is the entire nation, the economy may well represent a particular audit class only, where the audit class is sorted on the basis of some taxpayer characteristic which is directly observable but difficult or impossible to be manipulated by taxpayers (such as zip code, occupation and age). Sanchez and Sobel (1993) and Erard and Feinstein (1994), among others, interpret their audit rules within, and not across, audit classes. The same kind of interpretation is equally applicable here. This interpretation somewhat defuses the criticism that only the highest income taxpayers opt for FAT in our model. This interpretation also suggests that sorting taxpayers into as many different audit classes as possible could maximize the potential gain from introducing FATOTA.

Secondly, in addition to the possible benefits to taxpayers, the advantage of introducing FATOTA to the tax authority itself should not be overlooked. Running the regular tax system involves costly administration. For example, the IRS spent $7.6 billion in 1995 enforcing all varieties of federal taxation (see OMB (1995)). As observed by Chu (1990), administrative costs associated with tax auditing can be saved for the taxpayers choosing to pay FAT. This saving in administrative costs has not been taken into account yet, and will surely enforce the welfare-improving property of FATOTA. According to Bird and Casanegra de Jantscher (1992), a common constraint usually faced by tax reforms in developing countries is the scarcity of resources for tax administration. They even conclude that tax reforms requiring substantial additional administrative resources are usually doomed to failure, since the resources needed are unlikely to materialize fully or in a timely fashion. In view of this conclusion, the introduction of FATOTA as a complement to the regular tax system deserves serious consideration. Like plea bargaining in legal procedures which economizes on prosecution costs, the FATOTA system allows for a better allocation of administrative resources and hence could occupy importance in countries where the tax administration is overloaded.

Acknowledgements

We are grateful to a referee and an editor of this journal for their valuable suggestions and comments. Any remaining errors are our own.

References

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

Alm, J., 1988. Uncertainty tax policies, individual behavior, and welfare. American Economic Review 78, 237–245.

Andreoni, J., 1991. Reasonable doubt and the optimal magnitude of fines: should the penalty fit the crime? Rand Journal of Economics 22, 385–395.

Andreoni, J., 1992. IRS as loan shark: tax compliance with borrowing constraints. Journal of Public Economics 49, 35–46.

Andreoni, J., Erard, B., Feinstein, J., 1998. Tax compliance. Journal of Economic Literature 36, 818–860.

Arrow, K.J., 1970. Essays in the Theory of Risk-Bearing. North-Holland, Amsterdam.

Bebchuk, L.A., Kaplow, L., 1993. Optimal sanctions and differences in individuals' likelihood of avoiding detection. International Review of Law and Economics 13, 217–224.

Becker, G., 1968. Crime and punishment: an economic approach. Journal of Political Economy 76, 169–217.

Bird, R.M., Casanegra de Jantscher, M., 1992. The reform of tax administration. In: Bird, R.M., Casanegra de Jantscher, M. (Eds.), Improving Tax Administration in Developing Countries. IMF, Washington, DC.

Blumenthal, M., Slemrod, J., 1992. The compliance cost of the US individual income tax system: a second look after tax reform. National Tax Journal 45, 185–202.

Chu, C.Y., 1990. Plea bargaining with the IRS. Journal of Public Economics 41, 319–333.

Cremer, H., Gahvari, F., 1995. Tax evasion and the optimum general income tax. Journal of Public Economics 60, 235–249.

Erard, B., 1993. Taxation with representation: an analysis of the role of tax practitioners in tax compliance. Journal of Public Economics 52, 163–197.

Erard, B., Feinstein, J., 1994. Honesty and evasion in the tax compliance game. Rand Journal of Economics 25, 1–19.

Feldstein, M., 1976. On the theory of tax reform. Journal of Public Economics 6, 77–104.

Franzoni, L.A., 1999. Negotiated enforcement and credible deterrent. Economic Journal 109, 509–535.

Friedman, D., 1999. Why not hang them all: the virtue of inefficient punishment. Journal of Political Economy 107, S259–S269.

Gordon, J.P.F., 1989. Individual morality and reputation costs as deterrent to tax evasion. European Economic Review 33, 797–805.

Hall, A.P., 1995. Compliance costs of alternative tax systems: ways and means testimony. In: Tax Foundation Special Brief.

Klepper, S., Mazur, M., Nagin, D., 1991. Expert intermediaries and legal compliance: the case of tax preparers. Journal of Law and Economics 34, 205–229.

Kolm, S.C., 1973. A note on optimum tax evasion. Journal of Public Economics 2, 265–270.

Macho-Stadler, I., Perez-Castrillo, D., 1997. Optimal auditing with heterogeneous income sources. International Economic Review 38, 951–968.

Mirrlees, J., 1971. An exploration in the theory of optimum income taxation. Review of Economic Studies 48, 175–208.

Myles, G.D., Naylor, R.A., 1996. A model of tax evasion with group conformity and social customs. European Journal of Political Economy 12, 49–66.

OMB (Office of Management and Budget), 1995. The Budget of the United States Government, Fiscal Year 1996: Analytical Perspectives. GPO, Washington, DC.

Pestieau, P., Possen, U.M., Slutsky, S.M., 1994. The penalty for tax evasion when taxes are set optimally. CORE, Discussion paper 9416.

Reinganum, J., Wilde, L., 1985. Income tax compliance in a principal-agent framework. Journal of Public Economics 26, 1–18.

Roth, J.A., Scholz, J.T., Witte, A.D., 1989. Taxpayer Compliance, Vol. 1: An Agenda for Research, University of Pennsylvania Press, Philadelphia.

Sanchez, I., Sobel, J., 1993. Hierarchical design and enforcement of income tax policies. Journal of Public Economics 50, 345–369.

Sandmo, A., 1981. Income tax evasion, labor supply, and the equity-efficiency tradeoff. Journal of Public Economics 16, 265–288.

Schroyen, F., 1997. Pareto efficient income taxation under costly monitoring. Journal of Public Economics 65, 343–366.

Scotchmer, S., 1989. Who profits from taxpayer confusion? Economic Letters 29, 49–55.

Scotchmer, S., Slemrod, J., 1989. Randomness in tax enforcement. Journal of Public Economics 38, 17–32.

Seade, J.K., 1977. On the shape of optimal tax schedules. Journal of Public Economics 7, 203–235.

Slemrod, S., Bakaija, J., 1996. Taxing Ourselves. MIT Press, Cambridge.

Slemrod, S., Sorum, N., 1984. The compliance cost of the US individual income tax system. National Tax Journal 37, 461–474.

Ueng, K.L., Yang, C.C., 2000. Taxation with little administration. Journal of Public Economics 75, 145–156.

Watson, H., 1985. Tax evasion and labor markets. Journal of Public Economics 27, 231–246.

Yang, C.C., 1993. Optimal linear income tax with random revenue. Journal of Public Economics 52, 391–401.

Yitzhaki, S., 1974. A note on income tax evasion: a theoretical analysis. Journal of Public Economics 3, 201–202.

Yitzhaki, S., 1987. On the excess burden of tax evasion. Public Finance Quarterly 15, 123–137.