

A SIMPLE MODEL OF OPTIMAL TAX SYSTEMS: TAXATION, MEASUREMENT AND UNCERTAINTY*

by
SANJIT DHAMI
and
ALI AL-NOWAIHI†
University of Leicester

The neglect of administrative issues is a serious limitation of optimal tax theory, with implications for its practical applicability. We focus on an important class of administrative problems, namely that the tax bases are measured with some error. We also consider the full set of tax instruments. We find that consumption taxes can perform the ‘social insurance role of taxation’, a role previously ascribed only to income taxes. A combination of income and consumption taxes can hedge income and measurement-error risks better, relative to the imposition of either type of tax alone. The optimal tax rate is increasing in the precision with which the corresponding tax base is measured. The taxpayer engages in precautionary savings in response to income uncertainty and measurement problems. Differential commodity taxes tailored to the measurability characteristics of the different tax bases dominate uniform commodity taxes. However, as an economy becomes large, optimal taxes converge to uniform (or flat rate) taxes.

As economists have been aware, the omitted constraints on communication, calculation and administration of an economy . . . limit the direct applicability of the implications of this theory to policy problems. . . .’
(Diamond and Mirrlees, 1971)

1 INTRODUCTION

Two important considerations motivate this paper. First, despite their crucial importance in actual policy making, administrative issues are typically ignored in tax theory. Second, the practice of ignoring the full set of tax instruments, especially under uncertainty, leads to misleading results. We consider a model that is simple, tractable, and provides closed form solutions. It is perhaps instructive to look more closely at the two considerations that motivate the paper.

* Manuscript received 30.6.05; final version received 19.5.06.

† We are grateful to Joel Slemrod, Daniel Seidmann, Craig Brett, Rabah Amir and the referee for helpful comments and discussions. The usual disclaimers apply.

1.1 The Importance of Administrative Issues

In normative tax theory, issues of taxation and proposals for tax reform are typically evaluated on the basis of efficiency and equity considerations.¹ Administrative issues are ignored despite their importance in the actual implementation of tax policies.

Despite the neglect of administrative problems in the theory of taxation, general commentaries on tax policy often pose the choice between taxes, such as that between consumption and income taxes, in terms of the relative difficulty of measuring the two tax bases. Possibly the most interesting consequence of 'administrative issues' seems to be that tax bases are measured with some error. For instance, Devereux (1996, p. 14) writes, 'it is on administrative grounds that the proponents of the expenditure tax have the strongest case. This has largely to do with the problems of implementing a truly comprehensive income tax.' Bradford (1980) is very explicit: 'From this perspective, the winner of the great debate over the relative merits of the consumption versus the income tax rests on an issue of measurability'.

In the context of income taxation, Boadway and Wildasin (1996, p. 98) point to severe problems in the measurement of 'capital income'. They write: 'In principle this should include all forms of return to assets including interest, dividends, accrued capital gains, capital income from unincorporated business, imputed rent on consumer durables (especially housing) and the imputed return of assets such as transactions balances and insurance. These should all be indexed for inflation and should include an appropriate risk premium. Unfortunately the measurement of these items is difficult or impractical.' Mintz (1997, pp. 467–468) lists several problems in the measurement of a consumption tax base, both in its value-added tax (VAT) version and in its registered versus non-registered asset treatment. These problems include identification of taxpayers, issues of consumption versus business expenses, real versus financial transactions, wage versus self-employed income, treatment of losses, tracking of transactions etc.

The measurement of several tax bases can be especially difficult for developing countries. Burgess and Stern (1993, pp. 798–799) identify some of the relevant factors: insufficient staff with the appropriate skills, equipment, motivation or honesty; complex legal and tax structures; poor and inconsistent records that are often under the control of different tax authorities; lack of incentive-based remuneration etc. Although rarely acknowledged, developed countries often face similar problems. Fortin (1995, p. 2) writes: 'A substantial portion of Revenue Canada employees fails elementary tests of the knowledge of the tax system. Even our best experts admit that they find it very hard to keep up.'

¹Textbook treatments can be found in Atkinson and Stiglitz (1980) and Myles (1995).

Surveys of optimal taxation generally point to the lack of real world applicability of tax theory on administrative grounds. Heady (1996, p. 33) writes that ‘One way in which many models are unrealistic has already been mentioned: their neglect of administrative costs . . .’. Burgess and Stern (1993, p. 798) make similar remarks in the context of developing countries. Slemrod (1990, p. 157) provides a cogent overview of the issues and writes that ‘Differences in the ease of administering various taxes have been and will continue to be a critical determinant of appropriate tax policy’. Slemrod advocates incorporating administrative issues into ‘optimal tax theory’ to generate a unified ‘theory of optimal tax systems’. This paper can be viewed as one attempt in that direction.

1.2 The Full Set of Tax Instruments and Uncertainty

Under certainty, Atkinson and Stiglitz (1976), Atkinson (1977) and Deaton and Stern (1986) have shown that under some conditions commodity taxes are redundant in the presence of income taxes. However, under uncertainty, considering only a subset of the available taxes often results in erroneous conclusions. Nevertheless, the full set of tax instruments are generally omitted in optimal tax models involving uncertainty, as they were in the seminal papers by Eaton and Rosen (1980) and Varian (1980). The essential contribution of these papers was to identify the ‘social insurance role of taxation’ in the presence of income uncertainty.²

Cremer and Gahvari (1995, 1999) and Mirrlees (1990) provide further insights on optimal taxation in the presence of uncertainty.³ Cremer and Gahvari uncover a novel role for consumption taxes by distinguishing between goods that are consumed prior and posterior to the resolution of income uncertainty. However, their focus is not on administrative issues. Mirrlees (1990) comes closest to providing a theory of optimal tax systems under uncertainty but assumes that, while the income tax base is observed with some error, there are no such problems with measuring the consumption tax base.

This paper extends the basic model in Varian (1980) to take account of measurement problems with income as well as consumption tax bases and the full set of tax instruments is considered. The results are not tied to any particular source of measurement problems; rather such problems are taken as given. Attention is focused on ‘pure’ consumption and income taxes,

² For extensions, but mostly with a partial set of tax instruments, see the survey in Myles (1995).

³ Cremer and Gahvari (1995, 1999) and Mirrlees (1990) consider consumption taxes but omit income taxes. The argument is that ‘Given that the full set of commodity taxes is being used, the tax rate on the wage is superfluous and can be set equal to zero without imposing any restrictions’; see, for instance, Cremer and Gahvari (1995, p. 298). However, this does not hold true when both tax bases are observed with some measurement error. In that case (see below) the optimal income and consumption taxes will typically be strictly positive.

rather than on the specific institutional detail of any particular tax system. However, the model seems reasonably amenable to such extensions.⁴

The results are as follows. In the absence of income uncertainty and administrative problems, a poll tax is optimal. Under income uncertainty, income and consumption taxes perform a social insurance role. That such a role is not the exclusive domain of income taxes is not always reflected in the existing literature. When tax administration issues are taken into account, strictly positive income and consumption taxes are often optimal. Some combination of these taxes typically provides superior hedging of income and measurement error risk for the taxpayer. This result suggests a role that is similar to the idea of yardstick competition in the moral hazard literature. That the optimal consumption tax should be positive in the presence of an income tax is an important result given the originally pessimistic role for indirect taxes in Atkinson and Stiglitz (1976). In this respect the results in the paper also contribute to a growing literature that justifies a role for indirect taxes.⁵

Measurement error in a tax base reduces the optimal tax on that base. It also has 'spillover effects' on taxes levied on other tax bases. The relative magnitude of any two taxes is inversely related to the relative difficulty of measuring the respective tax bases. The magnitude of any tax is an increasing function of its social insurance role relative to the measurement error risk that it imposes. The taxpayer engages in precautionary savings, in response to uncertainty arising on account of income and tax administration. Differential commodity taxes, tailored to the measurability characteristics of the different tax bases, dominate uniform commodity taxes. However, as an economy becomes large, optimal tax rates converge to uniform tax rates.

In addition to providing sharp closed form results that have pedagogical merit, the main attractiveness of the model is its simplicity and tractability in dealing with fairly vexed questions. The questions posed in this paper are hardly novel, as the discussion above shows. However, one suspects that the lack of theoretical progress in the area owes much to lack of formal models that could provide a useful template for research. This paper is an attempt to fill that gap and hopefully provide the basis of a simple off-the-shelf model that can be used in a range of optimal tax situations under uncertainty.

Sections 2 through 4 adapt the model in Varian (1980) to the full set of tax instruments and measurement problems. Section 4.3 explores the implications for precautionary savings. Section 5 examines the issue of uniform versus differentiated taxes. Section 6 concludes.

⁴ See, for instance, the application in Dhimi (2002) to the 'registered' and the 'non-registered' asset treatment in consumption taxes.

⁵ For instance, Boadway *et al.* (1994) deal with tax evasion features; Cremer and Gahvari (1999) distinguish between those goods that are consumed before and after the resolution of income uncertainty. Cremer *et al.* (2001) introduce multidimensional heterogeneity among individuals. Each of these amendments to the Atkinson–Stiglitz framework uncovers a role for consumption taxes.

2 A MODEL

Consider the following two-period model as in Varian (1980). A representative taxpayer, when young, allocates first-period income, I_1 , between first-period consumption, C_1 , and saving, S . In the second period, the taxpayer is old and second-period income is $I_2 = S + \eta$, where η is a normally distributed random shock with mean zero and variance σ_η^2 . Varian (1980) interprets η as ‘idiosyncratic uncertainty’. Let second-period consumption be C_2 . Due to administrative problems, the government observes the imperfect signals, $I_2^O = I_2 + \varepsilon_I$ and $C_2^O = C_2 + \varepsilon_C$, of income and consumption, respectively, where ε_I and ε_C are measurement errors.⁶ These measurement errors are independent of the idiosyncratic uncertainty term η and are jointly normally distributed with zero mean, variances σ_I^2 and σ_C^2 and covariance σ_{IC} . Since, in practice, the two tax bases are measured with different methods,⁷ the distributions of ε_I and ε_C are likely to be different. Thus,

$$C_1 = I_1 - S \tag{1}$$

$$I_2 = S + \eta \tag{2}$$

$$I_2^O = I_2 + \varepsilon_I = S + \eta + \varepsilon_I \tag{3}$$

$$C_2^O = C_2 + \varepsilon_C \tag{4}$$

Taxes are levied only in the second period. First, the government levies an income tax on observed second-period income, I_2^O . The tax has a constant marginal rate θ . The government also makes a lump-sum payment T to the taxpayer ($T > 0$ denotes a transfer payment while $T < 0$ signifies a poll tax). After deducting the income tax, the taxpayer’s second-period disposable income is $I_2^D = I_2 - \theta I_2^O + T$. Next, the government levies a consumption tax on observed second-period consumption, C_2^O , at the constant marginal rate τ . Thus, the consumer’s second-period budget constraint is $C_2 + \tau C_2^O = I_2^D$. We get

$$I_2^D = I_2 - \theta I_2^O + T = S + \eta - \theta(S + \eta + \varepsilon_I) + T = (1 - \theta)(S + \eta) - \theta\varepsilon_I + T \tag{5}$$

$$I_2^D = C_2 + \tau C_2^O = C_2 + \tau(C_2 + \varepsilon_C) = (1 + \tau)C_2 + \tau\varepsilon_C \tag{6}$$

⁶ Measurement problems could arise from a wide variety of sources discussed in the introduction. Our model looks at a fairly generic problem; however, that does not specify the precise source of the measurement problem.

⁷ The issues are succinctly summarized in Boadway and Wildasin (1996, pp. 98–99). In a consumption tax, relative to an income tax, ‘it is no longer imperative to measure capital income on an accrual basis or to index capital income for the effect of inflation on asset values. Thus all accounting can be done on a cash flow basis which is relatively easier to administer. Furthermore, unlike in a comprehensive income tax, returns to capital which take an imputed form, such as rent on housing, need not be measured.’

From (5) and (6) we get

$$C_2 = \frac{1-\theta}{1+\tau}(S+\eta) - \frac{\theta}{1+\tau}\varepsilon_I - \frac{\tau}{1+\tau}\varepsilon_C + \frac{T}{1+\tau} \quad (7)$$

The government has an exogenous revenue requirement equal to R . Denoting by E the expectation operator with respect to the joint distribution of ε_I , ε_C and η , the government budget constraint is given by⁸

$$E(\theta I_2^0 + \tau C_2^0) - T = R \quad (8)$$

Hence, the government budget constraint is assumed to hold on average.⁹

Since $E\eta = E\varepsilon_I = E\varepsilon_C = 0$, substitution of (3), (4) and (7) into (8) gives

$$\frac{T}{1+\tau} = \frac{\theta+\tau}{1+\tau}S - R \quad (9)$$

Hence (7) and (9) give

$$C_2 = S - R + \frac{1-\theta}{1+\tau}\eta - \frac{\theta}{1+\tau}\varepsilon_I - \frac{\tau}{1+\tau}\varepsilon_C \quad (10)$$

The taxpayer's preference is of the constant absolute risk aversion (CARA) form, and is additively separable in C_1 and C_2 . Thus, expected utility is given by

$$E(U) = E(-e^{-\rho C_1}) + E(-e^{-\rho C_2}) \quad (11)$$

where $\rho > 0$ is the coefficient of absolute risk aversion. The strategy of using normally distributed error terms and CARA preferences is adopted from a related literature in agency theory.¹⁰ This substantially simplifies the problem and allows the derivation of closed form solutions.¹¹ From (1),

⁸ Cremer and Gahvari (1995, 1999) and Mirrlees (1990) write the government budget constraint with $\theta = 0$. This is admissible in their model because one of the two taxes is redundant in the presence of the other. However, in the presence of administrative problems, where the two tax bases might be faced with different measurement problems, setting $\theta = 0$ is not admissible. Indeed, as we will demonstrate below, both optimal taxes are generally positive.

⁹ This is a standard assumption for the government budget constraint under uncertainty; for instance, Varian (1980), Mirrlees (1990) and Myles (1995). We are grateful to the referee for asking us to consider the law of large numbers in this regard. We address this issue again in Section 5.3.

¹⁰ In multidimensional moral hazard models, a principal observes several imperfect signals of an agent's effort (e.g. Holmstrom and Milgrom, 1990, 1991). Holmstrom and Milgrom also show that, for this case, the optimal (private) incentive scheme between the principal and an agent is linear in the observed signals. Hence, viewing taxation as a (social) contract between the government and the taxpayer, Holmstrom and Milgrom's results suggest that the linear taxes used in this paper do not involve loss in generality.

¹¹ The lack of a general utility function would not seem to be important for our qualitative results. In terms of preferences, we only require risk aversion for our results. The exact form of attitudes to risk would have been crucial if we were interested in issues such as how additional income (in terms of magnitude or proportion of income) is allocated among a risky and a riskless asset (see Arrow, 1971). However, that is not the focus of our work.

$$E(-e^{-\rho C_1}) = -e^{-\rho(I_1 - S)} \tag{12}$$

Let $\mu = E(C_2)$ and $\sigma^2 = \text{var}(C_2)$. Then, from (10),

$$\mu = E(C_2) = S - R \tag{13}$$

$$\sigma^2 = \frac{1}{(1 + \tau)^2} [(1 - \theta)^2 \sigma_\eta^2 + \theta^2 \sigma_I^2 + \tau^2 \sigma_C^2 + 2\theta\tau\sigma_{IC}] \tag{14}$$

The normality of the disturbance terms implies that C_2 is also normally distributed. By a standard theorem in statistics or from direct integration, we get

$$E(-e^{-\rho C_2}) = -e^{-\rho\mu + (1/2)\rho^2\sigma^2} \tag{15}$$

From (11)–(13) and (15) we get

$$E(U) = -e^{-\rho(I_1 - S)} - e^{-\rho(S - R) + (1/2)\rho^2\sigma^2} \tag{16}$$

Note, from (14), that σ^2 is independent of S . Hence

$$\frac{d}{dS}(EU) = -\rho e^{-\rho(I_1 - S)} + \rho e^{-\rho(S - R) + (1/2)\rho^2\sigma^2} \tag{17}$$

3 SOLUTION TO THE OPTIMAL TAX PROBLEM

The government acts as the Stackelberg leader and commits to a tax vector (T, θ, τ) . The taxpayer observes (T, θ, τ) and then chooses the optimal value, S^* , of first-period saving, S . Set

$$\frac{d}{dS}(EU) = 0$$

in (17) to get

$$S^* = \frac{1}{2}(I_1 + R) + \frac{1}{4}\rho\sigma^2 \tag{18}$$

The first term in (18) captures the *intertemporal consumption-smoothing* component of saving while the second term, which is more fully explored in Section 4.3, captures the *precautionary* component of saving.

Once optimal values, θ^*, τ^* , for θ, τ have been chosen by the government, the corresponding optimal value, T^* , for T then follows from (9) and (18):

$$T^* = (\theta^* + \tau^*) \left[\frac{1}{2}(I_1 + R) + \frac{1}{4}\rho\sigma^{*2} \right] - (1 + \tau^*)R \tag{19}$$

where σ^{*2} is the value of σ^2 , from (14), corresponding to θ^*, τ^* . The corresponding value of saving is

$$S^{**} = \frac{1}{2}(I_1 + R) + \frac{1}{4}\rho\sigma^{*2} \tag{20}$$

Substitute from (18) into (16) to get the representative consumer’s indirect utility function, $V(\theta, \tau)$:

$$V(\theta, \tau) = -2e^{(1/2)\rho(R-I_1)+(1/4)\rho^2\sigma^2} \tag{21}$$

We assume that the objective of the government is to maximize the representative consumer’s indirect utility, $V(\theta, \tau)$, by a suitable choice of (θ, τ) . Let $V^* = V(\theta^*, \tau^*)$ be the maximized value of $V(\theta, \tau)$.

Lemma 1: (a) The indirect utility, $V(\theta, \tau)$, of the representative consumer is maximized if, and only if, the variance, σ^2 , of second-period consumption, C_2 , is minimized.

(b) If x is an exogenous variable then

$$\frac{\partial V^*}{\partial x} = \left(\frac{\partial V}{\partial x}\right)_{\theta^*, \tau^*} \quad \text{and} \quad \frac{\partial \sigma^{*2}}{\partial x} = \left(\frac{\partial \sigma^2}{\partial x}\right)_{\theta^*, \tau^*}$$

Proof: Since $V(\theta, \tau) < 0$ and ρ, R and I_1 are constant we get, from (21), that $V(\theta, \tau)$ is maximized when σ^2 is minimized. This establishes part (a). Part (b) is a special case of the envelope theorem. ■

From (14),

$$\frac{\partial \sigma^2}{\partial \theta} = -\frac{2(1-\theta)\sigma_\eta^2}{(1+\tau)^2} + \frac{2\theta\sigma_I^2}{(1+\tau)^2} + \frac{2\tau\sigma_{IC}}{(1+\tau)^2} \tag{22}$$

$$\frac{\partial \sigma^2}{\partial \tau} = -\frac{2(1-\theta)^2\sigma_\eta^2}{(1+\tau)^3} - \frac{2\theta^2\sigma_I^2}{(1+\tau)^3} + \frac{2\theta(1-\tau)\sigma_{IC}}{(1+\tau)^3} + \frac{2\tau\sigma_C^2}{(1+\tau)^3} \tag{23}$$

The first term in (22) captures ‘income-risk sharing’ between the government and the taxpayer on account of the income tax. This is the ‘social insurance effect’ in Eaton and Rosen (1980) and Varian (1980). The second term is the increased risk to the taxpayer on account of measurement errors in income. This is analogous to various forms of the ‘measurement risk effects’ in Stern (1982), Mirrlees (1990) and Dharami (2002). The third term is the ‘covariance effect’; correlation in the two measurement errors affects the overall risk facing the taxpayer, and hence has an affect on the optimal tax rates. The precise effect depends on whether the overall risk increases or decreases; this is examined in more detail below. An interpretation similar to that of (22) applies to (23) except for the last term, which takes account of measurement errors in the consumption tax base.

Proposition 1 derives some useful limiting results that help to build subsequent intuition about the model. They are all simple consequences of Lemma 1 and (14).

Proposition 1

- (a) If income is certain and there are no tax administration problems (i.e. $\sigma_\eta^2 = \sigma_I^2 = \sigma_C^2 = \sigma_{IC} = 0$), then any feasible combination of tax instruments is optimal. In particular, a poll tax is optimal.
- (b) When income is uncertain ($\sigma_\eta^2 > 0$) but administration problems are absent ($\sigma_I^2 = \sigma_C^2 = \sigma_{IC} = 0$), then $\theta^* = 1$.
- (c) Assume income certainty ($\sigma_\eta^2 = 0$).
 - (ci) If $\sigma_I^2 > 0$ but $\sigma_C^2 = 0$, then $\theta^* = 0$. If $\sigma_I^2 = 0$ but $\sigma_C^2 > 0$, then $\tau^* = 0$.
 - (cii) Suppose $\sigma_I^2 > 0$ and $\sigma_C^2 > 0$. Let r be the correlation coefficient between ε_I and ε_C . If $-1 < r < 1$ then a poll tax is optimal ($\theta^* = \tau^* = 0$). If $r = \pm 1$ then there is a multiplicity of non-zero tax solutions.
- (d) If σ_I^2 and σ_C^2 (and, hence, also σ_{IC}) are bounded but $\sigma_\eta^2 \rightarrow \infty$ then $\theta^* \rightarrow 1$.
- (e) If σ_η^2 , σ_C^2 and σ_{IC} are bounded but $\sigma_I^2 \rightarrow \infty$ then $\theta^* \rightarrow 0$.
- (f) If σ_η^2 , σ_C^2 and σ_{IC} are bounded but $\sigma_C^2 \rightarrow \infty$ then $\tau^* \rightarrow 0$.

Proof

- (a) Putting $\sigma_\eta^2 = \sigma_I^2 = \sigma_C^2 = \sigma_{IC} = 0$ in (14) gives $\sigma^2 = 0$, which is a minimum for any value of the tax instruments.
- (b) Putting $\sigma_I^2 = \sigma_C^2 = \sigma_{IC} = 0$, $\sigma_\eta^2 > 0$ in (14) gives $\sigma^2 = [(1 - \theta)/(1 + \tau)]^2 \sigma_\eta^2$, which is a minimum at $\theta = 1$.
- (c) Putting $\sigma_\eta^2 = 0$ in (14) gives $\sigma^2 = (\theta^2 \sigma_I^2 + \tau^2 \sigma_C^2 + 2\theta\tau\sigma_{IC}) / (1 + \tau)^2$.
 - (ci) If $\sigma_I^2 > 0$ but $\sigma_C^2 = 0$, then $\sigma_{IC} = 0$ and $\sigma^2 = \theta^2 \sigma_I^2 / (1 + \tau)^2$, which is minimized when $\theta = 0$. If $\sigma_I^2 = 0$ but $\sigma_C^2 > 0$, then $\sigma_{IC} = 0$ and $\sigma^2 = \tau^2 \sigma_C^2 / (1 + \tau)^2$, which is minimized when $\tau = 0$.
 - (cii) If $-1 < r < 1$, then σ^2 is a minimum when $\theta = \tau = 0$. If $r = -1$, then $\sigma^2 = (\theta\sigma_I - \tau\sigma_C)^2 / (1 + \tau)^2$, which is a minimum when $\theta\sigma_I = \tau\sigma_C$. If $r = 1$, then $\sigma^2 = (\theta\sigma_I + \tau\sigma_C)^2 / (1 + \tau)^2$, which is a minimum when $\theta\sigma_I = -\tau\sigma_C$.

Parts (d), (e) and (f) are obvious from Lemma 1 and (14). ■

Part (a) of Proposition 1 is consistent with the optimality of a poll tax under certainty, a result that is derived in Atkinson and Stiglitz (1976), Atkinson (1977) and Deaton and Stern (1986).

Part (b) illustrates the ‘social insurance role of taxation’ as in Varian (1980) and Eaton and Rosen (1980), namely that income taxes allow the

government and the taxpayer to share risks. However, its implications go slightly further. First, in the absence of tax administration issues, an income tax is superior to a consumption tax in sharing risk. Second, Cremer and Gahvari (1995, p. 53) ask why, given that labour supply is exogenous in Varian (1980), is not the optimal tax a 100 per cent income tax. Part (b) establishes the correctness of this intuition.

According to part (c), in the absence of a social insurance role of taxation, if a tax is associated with measurement error risk then it is best not to use it. If all tax bases are subject to measurement error risk, then a poll tax is optimal, because it imposes no such risk on the representative taxpayer. In the special case where $r = \pm 1$, i.e. the measurement errors are perfectly correlated, then the measurement error risk imposed on the taxpayer through one tax can be completely offset through the other tax. If $r = -1$ then risks are completely offset by strictly positive taxes, while if $r = +1$ then a combination of a consumption tax and an income subsidy (or the reverse) is optimal.

According to part (d), when σ_η^2 is high, the effect of income uncertainty can be countered by increasing income tax. High income tax also increases the measurement error risk. However, when $\sigma_\eta^2 \rightarrow \infty$, considerations of 'social insurance' outweigh measurement risk effects.

Parts (e) and (f) imply that a tax should not be used if there are severe problems in measuring its base, even if it would provide an excellent social insurance role.

4 OPTIMAL TAX FORMULAE AND COMPARATIVE STATICS RESULTS

In this section, we derive explicit formulae for the optimal tax rates and give some comparative statics results. The two cases, $\sigma_{IC} = 0$ and $\sigma_{IC} \neq 0$, are treated separately. We start with the simpler case: $\sigma_{IC} = 0$.

4.1 Uncorrelated Measurement Errors ($\sigma_{IC} = 0$)

Lemma 2: When $\sigma_{IC} = 0$, the optimal values of income and consumption tax rates are given by

$$\theta^* = \left(1 + \frac{\sigma_i^2}{\sigma_\eta^2} \right)^{-1} \quad (24)$$

$$\tau^* = \left(\frac{\sigma_C^2}{\sigma_i^2} + \frac{\sigma_C^2}{\sigma_\eta^2} \right)^{-1} \quad (25)$$

Proof: To maximize indirect utility of the representative consumer, minimize the variance, σ^2 , of second-period consumption (Lemma 1). Hence, set $\partial \sigma^2 / \partial \theta = 0$. This, together with (22) and $\sigma_{IC} = 0$, gives (24). Next, set $\partial \sigma^2 / \partial \tau = 0$. This, together with (23), $\sigma_{IC} = 0$ and (24), gives (25). ■

Proposition 2: The optimal tax rates θ^* and τ^* are increasing in the extent of income uncertainty, σ_η^2 . The optimal income tax rate, θ^* , is decreasing with the imprecision, σ_I^2 , in measuring income. The optimal consumption tax rate, τ^* , is decreasing with the imprecision, σ_C^2 , in measuring consumption. Furthermore, measurement problems create ‘spillover effects’: τ^* is increasing in σ_I^2 while θ^* is unaffected by σ_C^2 .

Proof: Obvious from (24) and (25). ■

As income uncertainty (captured by σ_η^2) increases, both tax rates optimally perform a ‘social insurance role’. As the informativeness of an observed tax base decreases (i.e. σ_I^2 or σ_C^2 increases), that tax is optimally reduced to mitigate the measurement error risk facing the taxpayer. In different contexts, Stern (1982), Mirrlees (1990) and Dhami (2002) also find that measurement problems reduce the optimal tax rate. The ‘spillover effect’ from one tax base to the other is demonstrated by $\partial\tau^*/\partial\sigma_I^2 > 0$. Despite $\sigma_{IC} = 0$, administrative problems with one tax base affect the optimal tax rate on the other tax base. The intuition can be seen by an examination of the expression for the taxpayer’s net consumption in (10). An increase in τ reduces $\theta_{E_I}/(1 + \tau)$, which is the exposure to measurement error risk in the income tax base. Finally, given the sequential nature of the two taxes, with the income tax levied before the consumption tax, the income tax is unaffected by measurement errors with the consumption tax base.

Proposition 3: The relative optimal tax rate, θ^*/τ^* , is inversely related to the relative imprecision, σ_I^2/σ_C^2 , in measuring the two tax bases:

$$\frac{\theta^*}{\tau^*} = \frac{\sigma_C^2}{\sigma_I^2} \tag{26}$$

Proof: Divide (24) by (25) and simplify to get the result. ■

Proposition 3 formalizes an intuitive idea. *Ceteris paribus*, the debate on the relative magnitude of the income and the consumption taxes rests, at least partially, as Bradford (1980) argues, on the issue of measurability. Optimal taxes on relatively easy to measure tax bases are relatively high. The results in Propositions 2 and 3 can have interesting implications for the following issues in tax theory.

1. *Direct versus indirect taxes.* Why do developing countries, unlike developed countries, raise the bulk of their tax revenues through indirect taxes rather than direct taxes? It is often argued that the explanation lies in the relative difficulty of measuring income in developing countries for reasons such as the paucity of recorded transactions, corrupt

tax administration etc. See, for example, Burgess and Stern (1993). This conforms to the result in Proposition 3.

2. *Taxation of fixed factors.* Economic theory demonstrates that taxes on fixed factors (e.g. land and capital) are efficient. Indeed, in the presence of such taxes there is no need for other taxes. However, why do such taxes account for a relatively small proportion of actual governmental tax revenues? One possible explanation, consistent with the predictions of Proposition 3, lies in the relative difficulty of measuring fixed factors. Two such taxes are considered below.

2(a) *Land taxes.* Bird (1974, p. 223) contends that ‘. . . the administrative constraint on effective land tax administration is so severe in most developing countries today that virtually all the more refined fiscal devices beloved of theorists can and should be discarded for this reason alone’. Similar problems are raised in Newbery (1987) and Skinner (1996). Land quality, which is one of the crucial elements in the definition of the land tax base, is hard to measure and requires ascertaining the soil type and quality, rainfall, irrigation facilities etc. Proxies for the land tax base such as capital value assessment, value of the produce on land, site value etc. are riddled with similar measurement problems. See, for example, Bird (1974). Hence, usage of the land tax is extremely limited and has historically declined.

2(b) *Capital stock taxes.* A capital tax is levied directly on the capital stock by state or federal authorities in many countries such as the USA, Canada and Germany at fairly low tax rates ranging from 0.25 to 0.50 per cent, with generous exemptions. Although a tax on the stock of capital that a firm owns is non-distortionary, there exist well-known difficulties in the measurement of the capital stock justifying the low taxation (or even exemption) of the tax base.

3. *Time inconsistency issues.* In an often cited example in the time inconsistency literature, the government announces that new capital is tax exempt. But once the new capital is in place, the government can renege and impose a 100 per cent capital tax which is *ex post* non-distortionary. Propositions 2 and 3 suggest that if measurement problems associated with the fixed tax base are acute, it is not efficient to impose a high confiscatory tax, even if the government has the discretion to do so. This argument provides a possible ‘optimal tax’ supplement to reputation-based explanations for the absence of 100 per cent capital taxes.¹²

¹² This is not meant to trivialize the time inconsistency literature which is clearly important, but to suggest a range of factors or assets to which the problem might not fully apply.

4.2 Correlated Measurement Errors ($\sigma_{IC} \neq 0$)

Lemma 3: Let r be the correlation between the errors, ε_I and ε_C , in measuring income and consumption, respectively:

$$r = \frac{\sigma_{IC}}{(\sigma_I^2 \sigma_C^2)^{1/2}} \tag{27}$$

When $\sigma_{IC} \neq 0$, the optimal values of income and consumption tax rates are given by

$$\tau^* = \frac{\sigma_I^2 + \sigma_\eta^2}{2r^2 \sigma_I^2} - \left[\left(\frac{\sigma_I^2 + \sigma_\eta^2}{2r^2 \sigma_I^2} \right)^2 - \frac{\sigma_\eta^2}{r^2 \sigma_C^2} \right]^{1/2} > 0 \tag{28}$$

$$\theta^* = \frac{\sigma_\eta^2 - \sigma_{IC} \tau^*}{\sigma_\eta^2 + \sigma_I^2} \tag{29}$$

Proof: To maximize indirect utility of the representative consumer, minimize the variance, σ^2 , of second-period consumption (Lemma 1). Hence, set $\partial \sigma^2 / \partial \theta = 0$. This, together with (22), gives (29). Note that (29) reduces to (24) as $\sigma_{IC} \rightarrow 0$. Next, set $\partial \sigma^2 / \partial \tau = 0$. This, together with (23) and (29), gives a quadratic equation in τ . This equation has two roots (both positive), one of which is reported as (28). Use L'Hôpital's rule to check that (28) reduces to (25) as $r \rightarrow 0$ (as it should). The other root goes to infinity as $r \rightarrow 0$ and therefore should be discarded (alternatively, one could appeal to second-order conditions to select the correct root). ■

From (28) we see that, if $\sigma_{IC} \neq 0$, then the optimal consumption tax is always positive.

Correlation in the measurement errors (i.e. $r \neq 0$) can have an important effect on the overall risk facing the taxpayer. This can be seen easily from the expression for the representative consumer's indirect utility (21) or from the first-order conditions in (22) and (23). To focus on the effects of r , assume that all other exogenous variables are fixed.

Proposition 4: (a) $\partial \tau^* / \partial r \leq 0 \Leftrightarrow r \leq 0$. (b) $\partial \theta^* / \partial r < 0$. (c) $\theta^* \leq 0 \Leftrightarrow r \geq (\sigma_C^2 / \sigma_I^2)^{1/2}$. (d) $\partial V^* / \partial r \geq 0 \Leftrightarrow \partial \sigma^{*2} / \partial r \leq 0 \Leftrightarrow \theta^* \leq 0 \Leftrightarrow r \geq (\sigma_C^2 / \sigma_I^2)^{1/2}$.

Proof: Inspecting (28) we see that τ^* is a decreasing function of r^2 . Since $dr^2/dr = 2r$, it follows that τ^* is a decreasing function of r if $r > 0$ and an increasing function of r if $r < 0$. This establishes (a). From (27) and (29) we get

$$\frac{\partial \theta^*}{\partial r} = - \frac{(\sigma_i^2 \sigma_c^2)^{1/2}}{\sigma_\eta^2 + \sigma_i^2} \left(\tau^* + r \frac{\partial \tau^*}{\partial r} \right)$$

From (28) we have $\tau^* > 0$. From (a), $r(\partial \tau^* / \partial r) > 0$. Thus $\partial \theta^* / \partial r < 0$. This establishes (b). From (29) $\theta^* = 0 \Leftrightarrow \sigma_\eta^2 = \sigma_{IC} \tau^*$. Substituting from (27) and (28) gives $\theta^* = 0 \Leftrightarrow r = (\sigma_c^2 / \sigma_i^2)^{1/2}$. Part (c) follows from this and (b). By the envelope theorem, or direct calculation, $\partial \sigma^{*2} / \partial r = 2\theta^* \tau^* (\sigma_i^2 \sigma_c^2)^{1/2} / (1 + \tau^*)^2$. Part (d) then follows from (21), (28) and (c). ■

Part (c) of Proposition 4 shows that if $r > (\sigma_c^2 / \sigma_i^2)^{1/2}$ then $\theta^* < 0$, i.e. the income ‘tax’ is in fact an income supplement. The correlation term, r , performs a role similar to that of ‘yardstick competition’ in the moral hazard literature, whereby the observation of two correlated signals of an agent’s effort allows the principal to filter some of the risk facing a risk-averse agent and allows for improved incentives (e.g. Holmstrom, 1982; Holmstrom and Milgrom, 1990, 1991). Proposition 4 shows that a variant of these results applies to a social contract between a government and a taxpayer. One is looking at circumstances when an increase in r increases the taxpayer’s share of the cake (in the analogous agency situation this is the agent’s share of the surplus). Part (d) of Proposition 4 shows that this is the case when $r > (\sigma_c^2 / \sigma_i^2)^{1/2}$.

4.3 Precautionary Behaviour

The model brings out some simple but important implications for precautionary behaviour. When income is *ex ante* uncertain and taxpayers are risk averse, one would expect them to engage in precautionary savings. Varian (1980) uses quadratic preferences, and thus the zero third derivative precludes precautionary behaviour.¹³ Strawczynski (1998) identifies precautionary savings by performing simulation techniques on a log utility version of the Varian (1980) model.

Since the third derivative is strictly positive for CARA preferences, the taxpayer engages in precautionary savings. From (20) and Lemma 1(b),

$$\frac{\partial S^{**}}{\partial \sigma_\eta^2} = \frac{1}{4} \rho \frac{\partial \sigma^{*2}}{\partial \sigma_\eta^2} = \frac{1}{4} \rho \left(\frac{\partial \sigma^2}{\partial \sigma_\eta^2} \right)_{\theta^*, \tau^*} = \frac{1}{4} \rho \left(\frac{1 - \theta^*}{1 + \tau^*} \right)^2$$

$$\frac{\partial S^{**}}{\partial \sigma_i^2} = \frac{1}{4} \rho \frac{\partial \sigma^{*2}}{\partial \sigma_i^2} = \frac{1}{4} \rho \left(\frac{\partial \sigma^2}{\partial \sigma_i^2} \right)_{\theta^*, \tau^*} = \frac{1}{4} \rho \left(\frac{\theta^*}{1 + \tau^*} \right)^2$$

$$\frac{\partial S^{**}}{\partial \sigma_c^2} = \frac{1}{4} \rho \frac{\partial \sigma^{*2}}{\partial \sigma_c^2} = \frac{1}{4} \rho \left(\frac{\partial \sigma^2}{\partial \sigma_c^2} \right)_{\theta^*, \tau^*} = \frac{1}{4} \rho \left(\frac{\tau^*}{1 + \tau^*} \right)^2$$

¹³ Precautionary savings require a positive third derivative (see, for example, Leland, 1968).

All three partial derivatives are positive; therefore, the taxpayer engages in precautionary savings with respect to future uncertainty arising from (1) income and (2) tax administration problems. Within models of precautionary savings, the second effect is a relatively novel result. These results complement the results in Strawczynski (1998) and provide the theoretical counterpart to his simulation results.

5 OPTIMAL COMMODITY TAXATION: DIFFERENTIATED OR UNIFORM?

The Ramsey model derives optimal consumption taxes in a representative taxpayer setting when efficiency is the sole objective of taxation. If all consumption goods are symmetric in all respects (i.e. identical compensated elasticities) then uniform commodity taxation is optimal. Otherwise, uniform taxation is not optimal. The model of this section shows that, for two identical commodities, if the respective tax bases are measured with different degrees of imprecision, then uniform commodity taxation is not optimal. Indeed, if one of the two commodities is measured relatively more precisely (all else being equal) it will be taxed at a higher rate.¹⁴ Although uniform commodity taxation is not optimal, under certain conditions optimal commodity taxes will converge to uniform taxes as the economy becomes ‘large’.

To fix ideas in a simple manner, consider a single-period model¹⁵ with n (possibly differentiated) goods. Preferences of the representative consumer take the CARA form and are additively separable over goods:

$$U = -\sum_{i=1}^n a_i e^{-\rho_i c_i} \quad \rho_i > 0 \quad a_i > 0 \quad \sum_{i=1}^n a_i = 1 \tag{30}$$

where c_i is the ‘real’ consumption of the i th good. The before-tax price of good i is $p_i > 0$ and is assumed exogenous. Hence, the pre-tax nominal expenditure on good i is $C_i = p_i c_i$. Rewrite (30) in terms of pre-tax expenditures, where $b_i = \rho_i / p_i$:

$$U = -\sum_{i=1}^n a_i e^{-\rho_i c_i} = -\sum_{i=1}^n a_i e^{-\rho_i C_i / p_i} = -\sum_{i=1}^n a_i e^{-b_i C_i} \tag{31}$$

The consumer has exogenously given income, $I + \eta$. The government observes this income with measurement error ϵ_0 ; hence, the observed income is $I^O = I + \eta + \epsilon_0$. The government also observes consumption with error; hence observed consumption is $C_i^O = C_i + \epsilon_i$, $i = 1, 2, \dots, n$, where ϵ_i is the

¹⁴ The results in this section can be easily modified to address the issue of uniform versus differentiated taxation of different sources of income. Since the treatment of these issues is analogous, but it is issues of uniform versus differentiated commodity taxes that typically receive more attention, it is omitted.

¹⁵ For a treatment of alternative consumption tax systems in a dynamic setting with consumption-savings choice under uncertainty and measurement problems, see Dhami (2002).

measurement error in the i th consumption tax base. The random variables $\eta, \epsilon_0, \epsilon_1, \epsilon_2, \dots, \epsilon_n$ are jointly normally distributed with zero means and covariance matrix

$$\Sigma = \begin{bmatrix} \sigma_{\eta\eta} & \sigma_{\eta 0} & \sigma_{\eta 1} & \sigma_{\eta 2} & \cdots & \sigma_{\eta n} \\ \sigma_{\eta 0} & \sigma_{00} & \sigma_{01} & \sigma_{02} & \cdots & \sigma_{0n} \\ \sigma_{\eta 1} & \sigma_{01} & \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{\eta 2} & \sigma_{02} & \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{\eta n} & \sigma_{0n} & \sigma_{1n} & \sigma_{2n} & \cdots & \sigma_{nn} \end{bmatrix} \tag{32}$$

The government levies income tax, on observed income $I^O = I + \eta + \epsilon_0$, at the constant marginal rate τ_0 . The government also levies a consumption tax on observed consumption of good $i, C_i^O = C_i + \epsilon_i$, at the constant marginal rate $\tau_i, i = 1, 2, \dots, n$. In addition, the government makes a lump-sum transfer, T , to the representative consumer ($T > 0$ is income support, $T < 0$ is a poll tax).

The sequence of moves is as follows. The government is the Stackelberg leader and chooses the tax vector $(T, \tau_0, \tau_1, \dots, \tau_n)$. The consumer observes his/her after-tax income, $I^D = I + \eta - \tau_0(I + \eta + \epsilon_0) + T$. He/she then allocates this disposable income to expenditure on the goods: $C_i + \tau_i C_i^O = C_i + \tau_i(C_i + \epsilon_i), i = 1, 2, \dots, n$. Thus, his/her budget constraint is

$$\sum_{i=1}^n (1 + \tau_i) C_i = (1 - \tau_0)(I + \eta) - \sum_{i=1}^n \tau_i \epsilon_i + T \tag{33}$$

The consumer maximizes his/her utility (31) subject to his/her budget constraint (33). This gives the consumer demand functions:

$$C_i = \frac{1}{b_i} \ln \frac{a_i b_i}{1 + \tau_i} + \frac{(1 - \tau_0)(I + \eta) - \sum_{j=0}^n \tau_j \epsilon_j - \sum_{j=1}^n (1 + \tau_j) \frac{1}{b_j} \ln \frac{a_j b_j}{1 + \tau_j} + T}{\sum_{j=1}^n (1 + \tau_j) \frac{1}{b_j}} \tag{34}$$

$$i = 1, 2, \dots, n$$

The government has an exogenous per capita revenue requirement, R , to which the per capita lump-sum transfer, T , is added. The government's total per capita tax revenue is $\tau_0 I^O + \sum_{i=1}^n \tau_i C_i^O = \tau_0(I + \eta + \epsilon_0) + \sum_{i=1}^n \tau_i(C_i + \epsilon_i)$. Assume that the government budget constraint holds in expected value per capita terms. Denote by E the expectation operator with respect to the joint distribution of $\eta, \epsilon_0, \epsilon_1, \epsilon_2, \dots, \epsilon_n$; then the government budget constraint is given by

$$\tau_0 I + \sum_{i=1}^n \tau_i EC_i = T + R \tag{35}$$

Substitute from (34) into (35) to get T :

$$T = t_0 I - R + \sum_{i=1}^n \frac{\tau_i}{b_i} \ln \frac{a_i b_i}{1 + \tau_i} + \frac{\sum_{j=1}^n \tau_j}{\sum_{j=1}^n \frac{1}{b_j}} \left(I - R - \sum_{i=1}^n \frac{1}{b_i} \ln \frac{a_i b_i}{1 + \tau_i} \right) \tag{36}$$

Use (36) to eliminate T from (34) to get

$$C_i^* = \frac{1}{b_i} \ln \frac{a_i b_i}{1 + \tau_i} + \frac{1}{b_i \sum_{j=1}^n \frac{1}{b_j}} \left(I - R - \sum_{j=1}^n \frac{1}{b_j} \ln \frac{a_j b_j}{1 + \tau_j} \right) + \frac{(1 - t_0) \eta - \sum_{j=0}^n \tau_j \varepsilon_j}{b_i \sum_{j=1}^n \frac{1 + \tau_j}{b_j}} \tag{37}$$

$i = 1, 2, \dots, n$

Let $\mu_i = EC_i^*$ and $s_i^2 = \text{var } C_i^*$ be respectively the mean and variance of C_i^* ; then

$$\mu_i = \frac{1}{b_i} \ln \frac{a_i b_i}{1 + \tau_i} + \frac{1}{b_i \sum_{j=1}^n \frac{1}{b_j}} \left(I - R - \sum_{j=1}^n \frac{1}{b_j} \ln \frac{a_j b_j}{1 + \tau_j} \right) \quad i = 1, 2, \dots, n \tag{38}$$

$$s_i^2 = \frac{(1 - \tau_0)^2 \sigma_{\eta\eta} + \sum_{j=0}^n \sum_{k=0}^n \tau_j \tau_k \sigma_{jk} - 2 \sum_{j=0}^n (1 - \tau_0) \tau_j \sigma_{nj}}{b_i^2 \left(\sum_{j=1}^n \frac{1 + \tau_j}{b_j} \right)^2} \tag{39}$$

$i = 1, 2, \dots, n$

From the above equations we get

$$\frac{\partial \mu_i}{\partial \tau_0} = 0 \quad i = 1, 2, \dots, n \tag{40}$$

$$\frac{\partial \mu_i}{\partial \tau_i} = \frac{1 - b_i \sum_{k=1}^n \frac{1}{b_k}}{(1 + \tau_i) b_i^2 \sum_{k=1}^n \frac{1}{b_k}} \quad i = 1, 2, \dots, n \tag{41}$$

$$\frac{\partial \mu_i}{\partial \tau_j} = \frac{1}{(1 + \tau_i) b_i b_j \sum_{k=1}^n \frac{1}{b_k}} \quad i, j = 1, 2, \dots, n \quad i \neq j \quad (42)$$

$$\frac{\partial s_i^2}{\partial \tau_0} = 2 \frac{\sum_{k=0}^n \tau_k (\sigma_{0k} + \sigma_{\eta k}) - (1 - t_0) (\sigma_{\eta\eta} + \sigma_{\eta 0})}{b_i^2 \left(\sum_{k=1}^n \frac{1 + \tau_k}{b_k} \right)^2} \quad i = 1, 2, \dots, n \quad (43)$$

$$\begin{aligned} \frac{\partial s_i^2}{\partial \tau_j} = & \frac{2}{b_i^2 \left(\sum_{k=1}^n \frac{1 + \tau_k}{b_k} \right)^3} \left\{ \left[\sum_{h=0}^n \tau_h \tau_{hj} - 2(1 - \tau_0) \sigma_{\eta j} \right] \left(\sum_{k=1}^n \frac{1 + \tau_k}{b_k} \right) \right. \\ & \left. - \frac{1}{b_j} \left[(1 - \tau_0)^2 \sigma_{\eta\eta} + \sum_{h=0}^n \sum_{k=0}^n \tau_h \tau_k \sigma_{hk} - 2 \sum_{k=0}^n (1 - \tau_0) \tau_k \sigma_{\eta k} \right] \right\} \quad (44) \end{aligned}$$

where $i, j = 1, 2, \dots, n$.

Analogous to (15), the expected value of the representative consumer’s indirect utility function is given by

$$V = - \sum_{i=1}^n a_i e^{-b_i \mu_i + (1/2) b_i^2 s_i^2} \quad (45)$$

We assume that the government chooses $\tau_0, \tau_1, \dots, \tau_n$ so as to maximize (45). The first-order conditions are

$$\frac{\partial V}{\partial \tau_j} = - \sum_{i=1}^n a_i e^{-b_i \mu_i + (1/2) b_i^2 s_i^2} \left(-b_i \frac{\partial \mu_i}{\partial \tau_j} + \frac{1}{2} b_i^2 \frac{\partial s_i^2}{\partial \tau_j} \right) = 0 \quad j = 0, 1, 2, \dots, n \quad (46)$$

Observe, from (39), (43) and (44), that $\frac{1}{2} b_i^2 s_i^2$ and $\frac{1}{2} b_i^2 (\partial s_i^2 / \partial \tau_j)$ are independent of i . Hence, i in $\frac{1}{2} b_i^2 s_i^2$ and $\frac{1}{2} b_i^2 (\partial s_i^2 / \partial \tau_j)$ can be taken to be any integer between 1 and n , say m . This allows (46) to be written as

$$\frac{\partial V}{\partial \tau_j} = \sum_{i=1}^n a_i b_i e^{-b_i \mu_i} \frac{\partial \mu_i}{\partial \tau_j} - \frac{1}{2} b_m^2 \frac{\partial s_m^2}{\partial \tau_j} \sum_{i=1}^n a_i e^{-b_i \mu_i} = 0 \quad j = 0, 1, 2, \dots, n \quad (47)$$

From (40) and (47) we get

$$\frac{\partial s_i^2}{\partial \tau_0} = 0 \quad i = 1, 2, \dots, n \quad (48)$$

From (43) and (48) we get

$$\tau_0^* = \frac{\sigma_{\eta\eta} + \sigma_{\eta 0} - \sum_{k=1}^n (\sigma_{0k} + \sigma_{\eta k}) \tau_k^*}{\sigma_{\eta\eta} + 2\sigma_{\eta 0} + \sigma_{00}} \tag{49}$$

which determines the optimal value of income tax, τ_0^* , in terms of optimal commodity tax rates, $\tau_1^*, \dots, \tau_n^*$.

Now we derive a useful formula from the other first-order conditions. Substituting from (38) into (47), and noting that

$$\frac{1}{\sum_{k=1}^n \frac{1}{b_k}} \left(I - R - \sum_{k=1}^n \frac{1}{b_k} \ln \frac{a_k b_k}{1 + \tau_k} \right)$$

is independent of i , we get

$$\frac{\partial V}{\partial \tau_j} = \left(\sum_{i=1}^n \frac{1 + \tau_i}{b_i} b_i \frac{\partial \mu_i}{\partial \tau_j} - \frac{1}{2} b_m^2 \frac{\partial s_m^2}{\partial \tau_j} \sum_{i=1}^n \frac{1 + \tau_i}{b_i} \right) \exp \left[\frac{1}{\sum_{k=1}^n \frac{1}{b_k}} \left(I - R - \sum_{k=1}^n \frac{1}{b_k} \ln \frac{a_k b_k}{1 + \tau_k} \right) \right] = 0$$

$$j = 0, 1, 2, \dots, n \tag{50}$$

Hence,

$$\sum_{i=1}^n \frac{1 + \tau_i}{b_i} b_i \frac{\partial \mu_i}{\partial \tau_j} - \frac{1}{2} b_m^2 \frac{\partial s_m^2}{\partial \tau_j} \sum_{i=1}^n \frac{1 + \tau_i}{b_i} = 0 \quad j = 0, 1, 2, \dots, n \tag{51}$$

Substitute from (41) and (42) into (51), and note that

$$\frac{1}{(1 + \tau_j) b_j \sum_{k=1}^n \frac{1}{b_k}}$$

is independent of i , to get

$$\left[\frac{1}{(1 + \tau_j) \sum_{k=1}^n \frac{1}{b_k}} - \frac{1}{2} b_j b_m^2 \frac{\partial s_m^2}{\partial \tau_j} \sum_{i=1}^n \frac{1 + \tau_i}{b_i} \right] = 1 \quad j = 1, 2, \dots, n \tag{52}$$

Change the summation index in (52) from i to k and rearrange, to get that the equilibrium values $\tau_0^*, \tau_1^*, \dots, \tau_n^*$ of the tax rates must satisfy

$$\frac{1}{1 + \tau_j^*} = \frac{1}{2} b_j b_m^2 \frac{\partial s_m^2}{\partial \tau_j} \sum_{k=1}^n \frac{1}{b_k} + \frac{\sum_{k=1}^n \frac{1}{b_k}}{\sum_{k=1}^n \frac{1 + \tau_k^*}{b_k}} \quad j = 1, 2, \dots, n \tag{53}$$

Equation (53) is useful for equilibrium analysis. In particular, from (44) and (53) we get

$$\frac{1}{1 + \tau_i^*} - \frac{1}{1 + \tau_j^*} = \left[\frac{b_i \sum_{h=0}^n \tau_h^* \sigma_{hi} - b_j \sum_{h=0}^n \tau_h^* \sigma_{hj}}{+ 2(1 - \tau_0^*)(b_j \sigma_{\eta j} - b_i \sigma_{\eta i})} \right] \frac{\sum_{k=1}^n \frac{1}{b_k}}{\left(\sum_{k=1}^n \frac{1 + \tau_k^*}{b_k} \right)^2} \tag{54}$$

$i, j = 1, 2, \dots, n$

Let $\sigma_i = (\sigma_{ii})^{1/2}$. Let $r_{\eta i}$ be the correlation between η and ε_i and r_{ij} the correlation between ε_i and ε_j ; $i, j = 0, 1, \dots, n$. Then $\sigma_{\eta i} = r_{\eta i} \sigma_{\eta} \sigma_i$ and $\sigma_{ij} = r_{ij} \sigma_i \sigma_j$. Thus, (54) becomes

$$\frac{1}{1 + \tau_i^*} - \frac{1}{1 + \tau_j^*} = \left[\frac{b_i \sum_{h=0}^n \tau_h^* r_{hi} \sigma_h \sigma_i - b_j \sum_{h=0}^n \tau_h^* r_{hj} \sigma_h \sigma_j}{+ 2(1 - \tau_0^*)(b_j r_{\eta j} \sigma_{\eta} \sigma_j - b_i r_{\eta i} \sigma_{\eta} \sigma_i)} \right] \frac{\sum_{k=1}^n \frac{1}{b_k}}{\left(\sum_{k=1}^n \frac{1 + \tau_k^*}{b_k} \right)^2} \tag{55}$$

$i, j = 1, 2, \dots, n$

5.1 Effect of Measurement Error on Tax Rates

Proposition 5: Consider the case where $\eta, \varepsilon_0, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are uncorrelated. Suppose $\tau_i^* > 0$. If $b_i \geq b_j$ and $\sigma_{ii} \geq \sigma_{jj}$ then, necessarily, $\tau_i^* < \tau_j^*$.

Proof: When $\eta, \varepsilon_0, \varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ are uncorrelated, (55) gives

$$\frac{1}{1 + \tau_i^*} - \frac{1}{1 + \tau_j^*} = \left(b_i \tau_i^* \sigma_{ii} - b_j \tau_j^* \sigma_{jj} \right) \frac{\sum_{k=1}^n \frac{1}{b_k}}{\left(\sum_{k=1}^n \frac{1 + \tau_k^*}{b_k} \right)^2} \quad i, j = 1, 2, \dots, n \tag{56}$$

Let $b_i \geq b_j$ and $\sigma_{ii} > \sigma_{jj}$. Suppose $\tau_i^* \geq \tau_j^*$. It then follows, from (56), that $1 + \tau_i^* < 1 + \tau_j^*$ and hence $\tau_i^* < \tau_j^*$, contrary to our supposition that $\tau_i^* \geq \tau_j^*$. Hence $\tau_i^* < \tau_j^*$. ■

In particular, if preferences over goods i and j are symmetric (so that $b_i = b_j$), except that the tax base of good j is measured more precisely, then it should be taxed at a higher rate.

5.2 Optimality of Uniform Taxation in a Large Economy

Consider a sequence of economies E_1, E_2, \dots , where economy E_n has n (possibly differentiated) goods, parameters $I(n), a_i(n), \rho_i(n), p_i(n), b_i(n) =$

$\rho_i(n)/p_i(n)$, error terms $\eta(n)$, $\varepsilon_i(n)$, covariance matrix $\Sigma(n)$, optimal tax rates $\tau_i^*(n)$, per capita revenue requirement $R(n)$ and per capita lump-sum transfer $T^*(n)$. Assume that there are numbers $\bar{\sigma}^2$, \bar{r} , \bar{b} and $\bar{\tau}$, independent of n , such that the following hold, for all n :

- (a) $\sigma_{\eta\eta}(n) < \bar{\sigma}^2$.
- (b) $\sigma_{ii}(n) < \bar{\sigma}^2$ for each $i = 0, 1, 2, \dots, n$.
- (c) $\sum_{j=1}^n |r_{ij}(n)| < \bar{r}$ for each $i = 0, 1, 2, \dots, n$.
- (d) $b_i(n) < \bar{b}$ for each $i = 1, 2, \dots, n$.
- (e) $|\tau_0^*(n)| < \bar{\tau}$.
- (f) $0 \leq \tau_i^*(n) < \bar{\tau}$ for each $i = 1, 2, \dots, n$.

Inequalities (a) and (b) put an upper bound, $\bar{\sigma}^2$, on all variances (contrast this with (d), (e) and (f) of Proposition 1). If we think of $|i - j|$ as a ‘distance’ between good i and good j , we suppose that, as this distance increases, the correlation between the respective measurement errors in their tax bases decreases. If, for example, $r_{ij}(n) = 1/2^{|i-j|}$, then (c) would be satisfied with $\bar{r} = 3$. If \bar{p} were an upper bound for all $\rho_i(n)$ and $\underline{p} > 0$ a lower bound for all prices, $p_i(n)$, then (d) would be satisfied with $\bar{b} = \bar{p}/\underline{p}$. (d) requires that we exclude sequences of economies where income taxes, or subsidies, become unbounded. (f) requires that we consider only sequences of economies where commodity taxes are non-negative and are all bounded above.

Proposition 6: Under assumptions (a)–(f), optimal commodity tax rates converge to a uniform tax rate.

Proof: To keep formulae to reasonable lengths, we shall not explicitly indicate the dependence of the parameters of economy E_n on n but, of course, they are dependent. From assumptions (a)–(f) we get

$$\left| b_i \sum_{h=0}^n \tau_h^* r_{hi} \sigma_h \sigma_i \right| \leq \bar{b} \bar{\sigma}^2 \bar{\tau} \sum_{h=0}^n |r_{hi}| \leq \bar{b} \bar{\sigma}^2 \bar{\tau} \bar{r} \tag{57}$$

$$\left| b_j \sum_{h=0}^n \tau_h^* r_{hj} \sigma_h \sigma_j \right| \leq \bar{b} \bar{\sigma}^2 \bar{\tau} \sum_{h=0}^n |r_{hj}| \leq \bar{b} \bar{\sigma}^2 \bar{\tau} \bar{r} \tag{58}$$

$$\left| 1 - \tau_0^* \right| \leq 1 + |\bar{\tau}| \tag{59}$$

$$|b_j r_{\eta j} \sigma_\eta \sigma_j| \leq \bar{b} \bar{\sigma}^2 \tag{60}$$

$$|b_i r_{\eta i} \sigma_\eta \sigma_i| \leq \bar{b} \bar{\sigma}^2 \tag{61}$$

$$\left| \frac{\sum_{k=1}^n \frac{1}{b_k}}{\left(\sum_{k=1}^n \frac{1+\tau_k^*}{b_k} \right)^2} \right| \leq \frac{\bar{b}}{n} \tag{62}$$

From (55)–(62) we get

$$|\tau_i^* - \tau_j^*| = |1 + \tau_i^*| |1 + \tau_j^*| \left| \frac{1}{1 + \tau_i^*} - \frac{1}{1 + \tau_j^*} \right| \leq \frac{2}{n} \bar{b}^2 \bar{\sigma}^2 (1 + \bar{\tau})^2 [\bar{\tau} + 2(1 + \bar{\tau})] \tag{63}$$

Hence, $|\tau_i^* - \tau_j^*| \rightarrow 0$ as $n \rightarrow \infty$. ■

Proposition 6 says that, as the economy becomes large, then optimal commodity taxes converge to uniform taxes, even with differentiated goods and measurement errors in tax bases. This provides theoretical support for the argument for uniform (or flat rate) taxes. This argument can also be extended to cover different sources of income. However, the argument does not exclude higher (or lower) rates of taxes levied to deal with several kinds of imperfections, for instance harmful externalities.

5.3 Volatility of Government Deficit

Here we investigate the effects of multiple sources of noise on the government budget constraint in a large economy. However, the situation is rather subtle in that we have to distinguish between an increase in the sources of noise and an increase in *total* noise. So, we modify assumption (b) of Section 5.2 to now read

(b') $\sigma_{00}(n) < \bar{\sigma}^2$ and $\sigma_{ii}(n) < (1/n^2) \bar{\sigma}^2$ for each $i = 1, 2, \dots, n$.

A rough interpretation of the new (b') is as follows. Suppose we increase the number of goods in the economy n -fold. To keep the total measurement error at the same order of magnitude, we replace each error term ε_i by ε_i/n . Hence its variance, σ_{ii} , is replaced by σ_{ii}/n^2 . We retain assumptions (a) and (c) and we weaken assumption (f) slightly. Thus, the new set of assumptions are

- (a) $\sigma_{\eta\eta}(n) < \bar{\sigma}^2$.
- (b') $\sigma_{00}(n) < \bar{\sigma}^2$ and $\sigma_{ii}(n) < (1/n^2) \bar{\sigma}^2$ for each $i = 1, 2, \dots, n$.
- (c) $\sum_{j=1}^n |r_{ij}(n)| < \bar{r}$ for each $i = 0, 1, 2, \dots, n$.
- (f') $|\tau_i^*(n)| < \bar{\tau}$ for each $i = 1, 2, \dots, n$.

Proposition 7: From assumptions (a), (b'), (c) and (f'), it follows that, as the number of goods increases, the government deficit converges in probability to zero.

Proof: Since the government's income from taxation is $\tau_0 I^O + \sum_{i=1}^n \tau_i C_i^O$ and its expenditure is $R + T$, it follows that its deficit is $D = \tau_0 I^O + \sum_{i=1}^n \tau_i C_i^O - R - T = \tau_0 (I + \eta + \varepsilon_0) + \sum_{i=1}^n \tau_i (C_i + \varepsilon_i) - R - T$. Substituting from the consumer's budget constraint (33) we get $D = I + \eta - \sum_{i=1}^n \tau_i \varepsilon_i - R$. The variance of this is $\text{var } D = \sum_{i=1}^n \sum_{j=1}^n \tau_i \tau_j \rho_{ij} \sigma_i \sigma_j$. From assumptions (a), (b') and (c)–(f) it follows that $\text{var } D \leq (1/n) \bar{r} \bar{\sigma}^2 \bar{\tau}^2$. Hence, $\text{var } D \rightarrow 0$ as $n \rightarrow \infty$. Since $E(D) = 0$ it follows, from Chebyshev's lemma, that D converges in probability to zero. ■

Thus, not only does the government budget constraint hold in expected value terms, but also the probability of it holding converges to one as the number of goods increases.

6 CONCLUSIONS

This paper incorporates administrative issues into optimal tax models under uncertainty, an endeavour termed by Slemrod (1990) a theory of 'optimal tax systems'. It focuses on one important implication of administrative problems, namely that tax bases are difficult or costly to measure. In spirit, the model is the one originally used in Varian (1980). However, it is extended to allow for tax administration considerations and the full set of (linear) tax instruments. One of the strengths of the model is its simplicity, and pedagogical merit in providing several closed form solutions. The model also provides a useful template for optimal tax theorists working in the general area of taxation under uncertainty, an area that has not progressed dramatically since the seminal works of Varian, Eaton and Rosen more than two decades ago.

Some of the results are as follows. Measurement problems reduce the optimal tax rates; in the limit as such problems become too severe, they might even override the 'social insurance role' of taxation. The social insurance role of taxation can be provided by consumption taxes, a role that in the literature is often ascribed to income tax alone. The relative magnitude of the income and consumption taxes is proportional to the ease of measuring the income tax base relative to the consumption tax base, a conjecture made by Bradford (1980). Errors in the measurement of a tax base can have 'spillover effects' by affecting the optimal tax rate on another base. In a stylized application of the basic model to consumption taxes, it is shown that, even in circumstances where the Ramsey taxes are predicted to be uniform, differences in the measurability characteristics of different commodities imply differentiated optimal commodity taxes. The model also derives implications for precautionary savings in the presence of income and administrative uncertainty.

Finally, although optimal tax rates are not uniform, they converge to uniform tax rates as the economy becomes large. In a large economy, this

suggests a *general principle of uniformity of taxes*, i.e. an approximately optimal tax system would involve, in the main, just three types of taxes: a uniform tax rate on most incomes, a uniform tax rate on most commodities and a lump-sum transfer. This excludes exceptions when a strong case can be made on other grounds such as externalities.

Although measurement problems could arise for a wide variety of reasons, these are considered exogenous in the paper. Endogenous treatment of measurement problems has the potential to produce a rich range of differentiated models of 'optimal tax systems'. This remains an important challenge and will reduce the gap between the theory and practice of taxation, an endeavour that was important to the pioneers of optimal tax theory.

REFERENCES

- Arrow, K. J. (1971). *Essays in the Theory of Risk Bearing*, Chicago, IL, Markham.
- Atkinson, A. B. (1977). 'Optimal Taxation and the Direct versus Indirect Tax Controversy', *Canadian Journal of Economics*, Vol. 10, pp. 590–606.
- Atkinson, A. B. and Stiglitz, J. E. (1976). 'The Design of Tax Structure: Direct versus Indirect Taxation', *Journal of Public Economics*, Vol. 6, pp. 55–75.
- Atkinson, A. B. and Stiglitz, J. E. (1980). *Lectures on Public Economics*, New York, McGraw-Hill.
- Bird, R. (1974). *Taxing Agricultural Land in Developing Countries*, Cambridge, MA, Harvard University Press.
- Boadway, R. and Wildasin, D. (1996). 'Taxation and Savings', in M. P. Devereux (ed.), *The Economics of Tax Policy*, Oxford, Oxford University Press.
- Boadway, R., Marchand, M. and Pestieau, P. (1994). 'Towards a Theory of the Direct–Indirect Tax Mix', *Journal of Public Economics*, Vol. 55, pp. 71–88.
- Bradford, D. (1980). 'The Case for a Personal Consumption Tax', in J. A. Pechman (ed.), *What Should Be Taxed: Income or Expenditure*, Washington, DC, Brookings Institute, pp. 75–113.
- Burgess, R. and Stern, N. (1993). 'Taxation and Development', *Journal of Economic Literature*, Vol. 31, pp. 762–830.
- Cremer, H. and Gahvari, F. (1995). 'Uncertainty and Optimal Commodity Taxation: in Defence of Commodity Taxes', *Journal of Public Economics*, Vol. 56, pp. 291–310.
- Cremer, H. and Gahvari, F. (1999). 'Uncertainty, Commitment and Optimal Taxation', *Journal of Public Economic Theory*, Vol. 1, pp. 51–70.
- Cremer, H., Pestieau, P. and Rochet, C. (2001). 'Direct versus Indirect Taxation: the Design of the Tax Structure Revisited', *International Economic Review*, Vol. 42, pp. 781–799.
- Deaton, A. and Stern, N. (1986). 'Optimally Uniform Commodity Taxes, Taste Differences and Lump Sum Grants', *Economics Letters*, Vol. 20, pp. 263–266.
- Devereux, M. P. (1996). 'Introduction', in M. P. Devereux (ed.), *The Economics of Tax Policy*, Oxford, Oxford University Press.
- Dhami, S. (2002). 'Optimal Consumption Taxes and Social Security under Tax Measurement Problems and Uncertainty', *International Tax and Public Finance*, Vol. 9, pp. 673–686.
- Diamond, P. A. and Mirrlees, J. A. (1971). 'Optimal Taxation and Public Production 1: Production Efficiency; 2: Tax Rules', *American Economic Review*, Vol. 61, pp. 8–27, 261–278.

- Eaton, J. and Rosen, H. S. (1980). 'Optimal Redistributive Taxation and Uncertainty', *Quarterly Journal of Economics*, Vol. 94, pp. 357–364.
- Fortin, P. (1995). 'An Efficient, Fair and Simple Income Tax for Canada', Invited Presentation to the Magna for Canada Scholarship Fund.
- Heady, C. (1996). 'Optimal Taxation as a Guide to Tax Policy', in M. P. Devereux (ed.), *The Economics of Tax Policy*, Oxford, Oxford University Press.
- Holmstrom, B. (1982). 'Moral Hazard in Teams', *Bell Journal of Economics*, Vol. 13, pp. 324–340.
- Holmstrom, B. and Milgrom, P. (1990). 'Regulating Trade among Agents', *Journal of Institutional and Theoretical Economics*, Vol. 146, pp. 85–105.
- Holmstrom, B. and Milgrom, P. (1991). 'Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design', *Journal of Law Economics and Organization*, Vol. 7, pp. 201–228.
- Leland, H. (1968). 'Saving and Uncertainty: the Precautionary Demand for Saving', *Quarterly Journal of Economics*, Vol. 82, pp. 465–473.
- Mintz, J. (1997). 'The Thorny Problem of Implementing New Consumption Taxes', *National Tax Journal*, Vol. 49, pp. 461–474.
- Mirrlees, J. A. (1990). 'Taxing Uncertain Incomes', *Oxford Economic Papers*, Vol. 42, pp. 34–45.
- Myles, G. (1995). *Public Economics*, Cambridge, Cambridge University Press.
- Newbery, D. (1987). 'Taxation and Development', in D. Newbery and N. Stern (eds), *The Theory of Taxation for Developing Countries*, New York, Oxford University Press.
- Skinner, J. (1996). 'If Agricultural Land Taxation is So Efficient, Why is it So Rarely Used?', in A. Braverman and J. E. Stiglitz (eds), *The Economics of Rural Organization: Theory, Practice, and Policy*, Oxford, Oxford University Press.
- Slemrod, J. (1990). 'Optimal Taxation and Optimal Tax Systems', *Journal of Economic Perspectives*, Vol. 4, pp. 157–178.
- Stern, N. (1982). 'Optimum Taxation with Errors in Administration', *Journal of Public Economics*, Vol. 17, pp. 181–212.
- Strawczynski, M. (1998). 'Social Insurance and the Optimum Piecewise Linear Income Tax', *Journal of Public Economics*, Vol. 69, pp. 371–388.
- Varian, H. R. (1980). 'Redistributive Taxation as Social Insurance', *Journal of Public Economics*, Vol. 14, pp. 49–68.