ELSEVIER Journal of Public Economics 56 (1995) 291–310

JOURNAL OF PUBLIC ECONOMICS

Uncertainty and optimal taxation: In defense of commodity taxes

Helmuth Cremer[a], Firouz Gahvari[b,*]

[a]GREMAQ, University of Toulouse, Toulouse, France
[b]University of Illinois at Urbana-Champaign, Champaign, IL 61820, USA and Virginia Polytechnic Institute & State University, Blacksburg, VA 24061, USA

Received January 1993, revised July 1993

Abstract

This paper re-examines the theory of optimal commodity taxation in the presence of a linear income tax, under wage uncertainty. There are two categories of goods: the consumption levels in one group are committed to before the resolution of uncertainty and those of the other after. The paper (i) characterizes the structure of the optimal commodity taxes in view of the insurance they provide against random wage movements, (ii) proves that optimal taxation requires a mix of differential commodity taxes and a uniform lump-sum tax, and (iii) demonstrates that the post-uncertainty goods should face a positive tax rate which is higher than the tax rate on the pre-committed goods.

Key words: Optimal; Uncertainty

JEL classification: H21; D80

1. Introduction

This paper re-examines the theory of optimal commodity taxation in the presence of a linear income tax, when the returns to labor are uncertain. In the process, it rediscovers a major role for commodity taxes as instruments

* Corresponding author.

SSDI 0047-2727(94)01422-K

for optimal tax policy – a role that has seriously been challenged in the literature (Atkinson and Stiglitz, 1976; Atkinson, 1977; Deaton, 1979; Deaton and Stern, 1986). The paper also sheds light on a controversial public policy debate, and puts forth a new argument in support of subsidizing housing. This is most interesting in the light of all the literature criticizing the implicit housing subsidies embedded in the US and UK tax systems. (For a survey of this literature, see Rosen, 1985).

The literature on optimal taxation has for the most part ignored the question of uncertainty. Given the large body of that literature, it is rather surprising to find so few papers dealing with this question.[1] Moreover, with one exception, none of these papers is concerned with the problem of optimal commodity taxation. The exception is Hamilton (1987) who shows, in a two-period model with wage uncertainty, that interest income taxation may enhance welfare by encouraging investment in human capital formation.[2] Our paper is another attempt to narrow this gap.

Eaton and Rosen's (1980a) pioneering work serves as the starting point of our study. It examines the structure of an optimal linear income tax in a community of identical consumers when there is uncertainty regarding the wage. In this setting, in the absence of uncertainty, the only tax instrument that would be utilized is the lump-sum element of the linear income tax. In the presence of uncertainty, on the other hand, a wage tax is also used as it reduces the riskiness of wage income. The wage tax thus works as an insurance mechanism ensuring a lesser tax payment in the event of a lower income.

As in Eaton and Rosen (1980a), we consider an economy consisting of identical consumers where the returns to labor are uncertain. We depart from Eaton and Rosen's framework, however, in two important ways. First, we consider a broader set of tax instruments. Rather than restricting the tax instruments to a linear income tax, we allow the government to also levy differential commodity taxes. Our setting includes the linear income tax as a special case which is obtained by taxing all commodities at a uniform rate. It therefore allows us to directly address the celebrated claim that, under certain restrictions on preferences, a linear income tax is a sufficient tool for

[1] These are: Eaton and Rosen (1980a,b,c), Varian (1980), Diamond et al. (1980), Persson (1983), Hamilton (1987), Koskela (1987), Richter (1987), Mazur (1989), Mirrlees (1990), Kaplow (1991) and Lundholm (1992). Stiglitz (1975), while not explicitly discussing optimal taxation, can also be given an interpretation along the lines of Eaton and Rosen (1980a).

[2] Varian (1980) also comes close. His is, in essence, a two-good model with exogenous labor supply and stochastic income where consumers are identical. He addresses the question of optimal taxation that arises because of the stochastic nature of income. However, the only tax instrument that he allows, in addition to a uniform lump-sum tax, is a tax on one good (second period consumption) only.

optimal taxation. The original claim is due to Atkinson and Stiglitz (1976). It has been generalized by Atkinson (1977), Deaton (1979) and Deaton and Stern (1986).

The second crucial difference between our framework and that of Eaton and Rosen lies in our recognition that consumption may take place *before* as well as after the resolution of uncertainty. Eaton and Rosen (1980a), and Mirrlees (1990), allow consumption to take place only *after* the resolution of uncertainty.[3] This is rather curious in the light of Mirrlees' own justification of the uncertainty framework. He writes "People deciding what labour to supply are uncertain what labour income their efforts will produce, particularly *when the results are delayed*, as with training and career choices" (p. 34, emphasis added). Surely, a delay must also mean some early consumption. Our model, like Eaton and Rosen's, is a single-period model. We thus short-circuit the timing problem by distinguishing among the commodities on the basis of commitment: The consumption levels of one category must be committed to before the resolution of the uncertainty and the other consumption levels after.

This particular way of looking at the problem distinguishes amongst commodities on the basis of the timing of consumption and not on the basis of their physical characteristics. However, in many instances the two characteristics are closely linked. Durable goods in general, and housing in particular, are consumed for many periods subsequent to their purchase. One often buys a house *prior* to the realization of an uncertain future income through a mortgage. This is also true of automobiles. In a world of uncertainty, given the rather high transaction costs involved in the resale of these goods, one may consider them as goods where consumption levels are committed to before the resolution of uncertainty.

Within this framework, we first characterize the structure of optimal commodity taxes on both categories of goods. We prove that optimal taxation requires a mix of differential commodity taxes and a uniform lump-sum tax. Intuitively, this result is due to the insurance that taxation of post-uncertainty (but not pre-committed) goods provides. As in Eaton and Rosen (1980a), the need for insurance implies that taxation is welfare enhancing. At the same time, commitment implies that one does not need insurance for *all* goods, as pre-committed goods require no insurance. Consequently, differential commodity taxes become useful in that they allow an individual to insure himself differently for different goods. The result is quite robust; it does not go away by imposing restrictions on preferences as

[3] Varian (1980), Diamond et al. (1980) and Hamilton (1987), on the other hand, have explicit two-period models in which first-period consumption takes place before, and the second-period consumption after, the resolution of uncertainty.

long as consumers are risk averse. Commitment in the presence of uncertainty thus reclaims the lost role of commodity taxes as instruments of optimal tax policy.[4]

We also prove that goods whose consumption levels are committed to *after* the resolution of the uncertainty (goods with the insurance property), should face a positive tax rate which is higher than the tax rate on goods in the other category. This result underscores the role that differential commodity taxation plays as a source of insurance. In the absence of insurance markets, a risk averse person concerned about the possibility of becoming 'poor' tends to 'under-consume' the goods to which he has to commit to, and to 'over-consume' the post-uncertainty goods (all relative to a situation where full insurance is available). A policy of taxing post-uncertainty goods induces the individual to consume less of these and more of the pre-committed goods (which everybody consumes equally). This has the effect of providing an individual with some measure of insurance as it reduces the gap between total after-tax expenditures in cases of bad and good states of the world. It also allows the individual to achieve this reduction by changing his consumption of different goods differently.[5]

2. The economy

The economy consists of a large number of identical individuals. Each person, when deciding on the allocation of his time between labor and leisure, faces uncertainty about the returns to his labor. The uncertain wage, w, is continuously distributed over some interval. The goods the person purchases and consumes fall into two categories: (i) goods with levels of consumption which must be committed to, along with labor supply, *prior* to the resolution of the uncertainty (purchases of housing and other consumer durables) and (ii) goods with levels of consumption which become known *after* the resolution of uncertainty (purchases of non-housing and other non-durable goods). There are m goods in the first category and n goods in the second, with m and $n \geqslant 1$. The vector of goods in category (i) is denoted by $\underline{y} \equiv (y_1, y_2, \ldots, y_m)$ and in category (ii) by $\underline{x} \equiv (x_1, x_2, \ldots, x_n)$. Labor supply is denoted by L and the time endowment is normalized to one.

[4] Of course, differential commodity taxation may be useful even without commitment. However, as shown by Atkinson and Stiglitz (1976), Atkinson (1977), Deaton (1979) and Deaton and Stern (1986), this form of taxation may be unnecessary with certain restrictions on preferences.

[5] This is, in spirt, similar to Richter's (1987) argument for rate differentiation in inheritance taxation. He argues that progression should be higher the larger the non-expected component of one's income (the higher the insurance motive).

2.1. The individual

Assume preferences are separable in $\underline{x}$, $\underline{y}$ and L, and are represented by

$$U = u(\underline{x}) + \psi(\underline{y}) + \varphi(1 - L) , \tag{1}$$

where U is twice continuously differentiable, (strictly) increasing in $\underline{x}$ and $\underline{y}$ and (strictly) decreasing in L. Moreover, to ensure risk aversion, assume that u is strictly concave. Denote the non-wage income by G, the price of x_k by p_k and the price of y_i by q_i.[6] The individual's budget constraint, which must hold for every w, is given by

$$\sum_{k=1}^{n} p_k x_k(w) + \sum_{i=1}^{m} q_i y_i = wL + G . \tag{2}$$

The individual's decision problem may be decomposed into two parts: He chooses L and $\underline{y}$ in the first stage and $\underline{x}$ in the second stage. Of course, in the first stage, the individual will take into account his (optimal) second stage decisions regarding $\underline{x}$.

Formally, the individual's optimization problem can be stated as follows. In the second stage, after the resolution of the uncertainty, the individual has

$$I \equiv wL + G - \sum_{i=1}^{m} q_i y_i \tag{3}$$

to spend on x_ks. He solves a standard consumer problem under certainty and chooses $\underline{x}$ to maximize $u(\underline{x})$ subject to having an income equal to I. This optimization problem yields the demand functions

$$x_k = \hat{x}_k(\underline{p}, I), k = 1, 2, \ldots, n , \tag{4}$$

where $\underline{p} \equiv (p_1, p_2, \ldots, p_n)$. Substituting from (4) into $u(\underline{x})$ gives the indirect utility function corresponding to u:

$$v = v(\underline{p}, I) \equiv u(\hat{x}_1(\underline{p}, I), \hat{x}_2(\underline{p}, I), \ldots, \hat{x}_n(\underline{p}, I)) . \tag{5}$$

Note that $\hat{x}_k$s and v are the standard (ordinary) demand and indirect utility functions possessing all the traditional properties. Specifically, we have

$$\frac{\partial v}{\partial I} = \alpha , \tag{6a}$$

$$\frac{\partial v}{\partial p_k} = -\alpha \hat{x}_k, k = 1, 2, \ldots, n , \tag{6b}$$

[6] Unless otherwise stated, k always runs from 1 to n and i from 1 to m.

where α denotes the marginal utility of income.

In the first stage, w and thus I, are random variables. We denote the probability distribution function of w by $F(w)$ and use E for the expectation operator. The problem of the individual is to choose $\underset{\sim}{y}$ and L to maximize

$$V = \mathrm{E}[v(\underset{\sim}{p}, I)] + \psi(\underset{\sim}{y}) + \varphi(1 - L)\,, \tag{7}$$

with I defined by (3). Assuming an interior solution, the first-order conditions are

$$\frac{\partial V}{\partial L} = \mathrm{E}(\alpha w) - \varphi'(1 - L) = 0\,, \tag{8a}$$

$$\frac{\partial V}{\partial y_i} = -q_i \mathrm{E}(\alpha) + \frac{\partial \psi(\underset{\sim}{y})}{\partial y_i} = 0\,, \quad i = 1, 2, \ldots, m\,. \tag{8b}$$

Eqs. (8a)–(8b) along with Eq. (3) determine the supply of L and the demand for y_is as functions of $\underset{\sim}{q}$, $\underset{\sim}{p}$ and G: $L = L(\underset{\sim}{q}, \underset{\sim}{p}, G)$ and $y_i = y_i(\underset{\sim}{q}, \underset{\sim}{p}, G)$. (Of course, the functional forms for L and y_i also depend on the probability distribution of w). Next, substituting for L and y_is in (3) yields $I = I(w, \underset{\sim}{q}, \underset{\sim}{p}, G)$, which is then substituted in (4) to give the demand functions for x_ks: $x_k = x_k(w, \underset{\sim}{q}, \underset{\sim}{p}, G)$.[7] Note that before the resolution of the uncertainty x_ks are random variables while L and y_is are not.

Finally, we can substitute the demand and supply functions into (7) to obtain the maximum value of V, denoted by V^*, as a function of $\underset{\sim}{q}$, $\underset{\sim}{p}$ and G

$$V^* = V^*(\underset{\sim}{q}, \underset{\sim}{p}, G)\,. \tag{9}$$

This function is a major ingredient in the optimal taxation problem considered below. It has several interesting properties which are easily derived using the envelope theorem; they are[8]

[7] Formally we have: $x_k(w, \underset{\sim}{q}, \underset{\sim}{p}, G) \equiv \hat{x}_k(\underset{\sim}{p}, I(w, \underset{\sim}{q}, \underset{\sim}{p}, G))$.

[8] The individual's problem can be formulated more succinctly as a one-stage problem of choosing $\underset{\sim}{x}(w)$, $\underset{\sim}{y}$ and L to maximize $\mathrm{E}[u(\underset{\sim}{x}(w)) + \psi(\underset{\sim}{y}) + \varphi(1 - L)]$. The first-order conditions of this problem include Eqs. (8a)–(8b). One can then

$$\mathrm{V}^* = \mathrm{V}^*(\underset{\sim}{q}, \underset{\sim}{p}, G) \equiv \max_{\underset{\sim}{x}(w), \underset{\sim}{y}, L} \Big\{ E[u(\underset{\sim}{x}(w))] + \psi(\underset{\sim}{y}) + \varphi(1 - L) \text{ such that } \sum_{k=1}^{n} p_k x_k(w) + \sum_{i=1}^{m} q_i y_i - wL - G = 0 \Big\}\,,$$

where its differentiation, upon application of the envelope theorem, yields equations (10a)–(10c).

We have opted for an explicitly two-stage statement of the problem, as this allows us later to prove lemma 1 and proposition 3 in quite a straightforward way.

$$\frac{\partial V^*}{\partial G} = \mathrm{E}(\alpha)\,, \tag{10a}$$

$$\frac{\partial V^*}{\partial q_i} = -y_i\,\mathrm{E}(\alpha)\,, \tag{10b}$$

$$\frac{\partial V^*}{\partial p_k} = -\mathrm{E}(\alpha x_k)\,. \tag{10c}$$

2.2. Compensated demand functions

The concept of compensated demand has a natural counterpart in our setting, which we use extensively below. However, in the presence of uncertainty, we must first define what we mean by a compensated demand function and study its properties, as they somewhat differ from those obtained under certainty.

Define the expenditure function

$$\mathbf{e}(\underline{q}, \underline{p}, V)\,, \tag{11}$$

as the minimum level of non-wage income for which the consumer can achieve a level of expected utility V if prices are $\underline{q}$ and $\underline{p}$. This function can be obtained by inverting the indirect utility function defined by Eq. (9) so that $\mathbf{e}(\underline{q}, \underline{p}, V^*) = G$.[9] The compensated demand functions are then defined as

$$L^c(\underline{q}, \bar{p}, V^*) \equiv L(\underline{q}, \underline{p}, \mathbf{e}(\underline{q}, \underline{p}, V^*))\,, \tag{12a}$$

$$y_i^c(\underline{q}, \underline{p}, V^*) \equiv y_i(\underline{q}, \underline{p}, \mathbf{e}(\underline{q}, \underline{p}, V^*))\,, \tag{12b}$$

$$x_k^c(w, \underline{q}, V^*) \equiv x_k(w, \underline{q}, \underline{p}, \mathbf{e}(\underline{q}, \underline{p}, V^*))\,. \tag{12c}$$

In words, we substitute the expenditure function into the ordinary demand functions to derive the corresponding compensated demands.[10]

Next, we derive the following relationships by totally differentiating (9) with respect to q_i and p_k, for a given value of V^*, and using (10a)–(10c) and (12a)–(12c)

[9] Since $\mathrm{E}(\alpha) > 0$, (10a) implies that V^* is strictly increasing in G.

[10] Alternatively, one can proceed as follows. By fixing V in (7), say at V^*, and substituting for I from (3), one may define G implicitly as a function of V^*, $\underline{q}$, $\underline{p}$, L and $\underline{y}$. Minimizing this function with respect to L and y_is gives the compensated demand functions for L and y_is. Substituting these values back in the implicit function for G gives the expenditure function. Derivations of x_k^cs follow immediately.

$$\frac{\partial \mathrm{e}}{\partial q_i} = y_i^c \,, \tag{13a}$$

$$\frac{\partial \mathrm{e}}{\partial p_k} = \frac{\mathrm{E}(\alpha x_k^c)}{\mathrm{E}(\alpha)} \,. \tag{13b}$$

Note that (13a) is identical to the standard relationship that is obtained in the absence of uncertainty. This, however, is not true for (13b) since $\mathrm{E}(\alpha x_k^c)$ is not, in general, equal to $\mathrm{E}(\alpha)\mathrm{E}(x_k^c)$ (see our discussion at the end of Section 3).

Finally, from (12a)–(12c) and (13a)–(13b) we can derive the following 'Slutsky type' decompositions[11]

$$\frac{\partial y_j^c}{\partial q_i} = \frac{\partial y_j}{\partial q_i} + y_i^c \frac{\partial y_j}{\partial G} \,, \tag{14a}$$

$$\frac{\partial y_j^c}{\partial p_k} = \frac{\partial y_j}{\partial p_k} + \frac{\mathrm{E}(\alpha x_k^c)}{\mathrm{E}(\alpha)} \frac{\partial y_j}{\partial G} \,, \tag{14b}$$

with similar expressions for x_s^c and L^c. Not surprisingly, the decompositions of derivatives with respect to p_k's differ from the standard expressions.

3. The government

The government's tax instruments consist of commodity taxes and a linear income tax with a lump-sum element. Denote the tax rate on y_i by τ_i and on x_k by t_k. Assume all producer prices are constant and normalized at one. Continue to use q_i and p_k to denote consumer prices so that we have $q_i = 1 + \tau_i$ and $p_k = 1 + t_k$. As to the income tax parameters, we use G to denote the lump–sum element of the linear tax system. Given that a full set of commodity taxes are being used, the tax rate on the wage is superfluous and can be set at zero without imposing any restrictions. The per-capita tax revenue requirement is R_0, where $R_0 \geqslant 0$. The government's problem is to set τ_is, t_ks and G to maximize the (indirect) utility of a representative individual, subject to a tax revenue constraint. More formally, it solves

$$\max_{\tau_i, t_k, G} V^*(\underset{\sim}{q}, \underset{\sim}{p}, G) \tag{15}$$

[11] Unless otherwise stated, j always runs from 1 to m and s from 1 to n.

subject to

$$\sum_{i=1}^{m} \tau_i y_i + \sum_{k=1}^{n} t_k \, \mathrm{E}(x_k) = G + R_0 \, , \tag{16}$$

where the population is assumed to be large enough for the realized means of the random variables to be equal to their expected values. Note that y_is and x_ks in (16) are given by the individual's demand functions which, as shown in Section 2, depend on $\underset{\sim}{p}$, $\underset{\sim}{q}$ and G.

Assuming an interior solution, the first-order conditions are

$$\frac{\partial V^*}{\partial G} + \lambda \left[\sum_i \tau_i \frac{\partial y_i}{\partial G} + \sum_k t_k \, \mathrm{E}\left(\frac{\partial x_k}{\partial G}\right) - 1 \right] = 0 \, , \tag{17a}$$

$$\frac{\partial V^*}{\partial \tau_j} + \lambda \left[y_j + \sum_i \tau_i \frac{\partial y_i}{\partial q_j} + \sum_k t_k \, \mathrm{E}\left(\frac{\partial x_k}{\partial q_j}\right) \right] = 0 \, , \qquad j = 1, 2, \dots, m \, , \tag{17b}$$

$$\frac{\partial V^*}{\partial t_s} + \lambda \left[\mathrm{E}(x_s) + \sum_i \tau_i \frac{\partial y_i}{\partial p_s} + \sum_k t_k \, \mathrm{E}\left(\frac{\partial x_k}{\partial p_s}\right) \right] = 0 \, , \qquad s = 1, 2, \dots, n \, , \tag{17c}$$

where λ is the Lagrangean multiplier associated with the government's budget constraint. The optimal tax rates may then be found by solving (17a) for λ and substituting in (17b)–(17c), making use of (10a)–(10c) and (14a)–(14b) in the resulting expression and simplifying. This is summarized in proposition 1 where cov(·, ·) denotes covariance.

Proposition 1. The optimal commodity taxes are characterized by

$$\sum_i \tau_i \frac{\partial y_i^c}{\partial q_j} + \sum_k t_k \, \mathrm{E}\left(\frac{\partial x_k^c}{\partial q_j}\right) = 0 \, , \tag{18a}$$

$$\sum_i \tau_i \frac{\partial y_i^c}{\partial p_s} + \sum_k t_k \, \mathrm{E}\left(\frac{\partial x_k^c}{\partial p_s}\right) = \frac{\mathrm{cov}(\alpha, x_s)}{\mathrm{E}(\alpha)} \, . \tag{18b}$$

The system of equations (18a)–(18b) determines the optimal tax rates τ_is and t_ks.[12] The equations show how the optimal tax rates strike the right balance between efficiency cost of taxation and the benefit of insurance in presence of commitment. Each left-hand side term in (18a) measures the change in welfare brought about by a change in τ_j taking into account the fact that the additional tax revenue affects the lump-sum term G.[13] Similarly, each left- minus right-hand side term in (18b) measures the net impact of a change in t_s.

The left-hand side of Eqs. (18a)–(18b) have a familiar flavor. Each one measures the change in welfare associated with the substitution effects created by an increase in one of the tax rates. This is similar to the excess burden terms in traditional optimal taxation models. Of course, not every single term necessarily reflects a cost (as opposed to a benefit). As is usual in a second-best setting, a tax increase in one of the commodities can very well reduce the existing distortions for the others.[14]

The right-hand side of (18b) includes a covariance term. Such terms are quite common in the optimal income tax literature and usually capture the redistributive benefits of taxation. In our setting, there is no redistribution per se since individuals are identical ex-ante and the government is concerned with expected utility. Still, because of the uncertainty, the tax on x_s provides some measure of insurance which is nothing else than (ex-post) redistribution among individuals. The benefit of such insurance is measured by $-\mathrm{cov}(\alpha, x_s) \equiv \mathrm{E}(\alpha)\mathrm{E}(x_s) - \mathrm{E}(\alpha x_s)$. To see this, note that the direct effect of an increase in t_s is, from (10c), to reduce expected utility by $\mathrm{E}(\alpha x_s)$. However, this also permits, from (16), an increase in G equal to $\mathrm{E}(x_s)$ and, thus, an increase of $\mathrm{E}(\alpha)\mathrm{E}(x_s)$ in expected utility. Moreover, with u being strictly concave, α is a decreasing function of income. Accordingly, the

[12] Assuming these equations are linearly independent.

[13] To be precise, it is equal to

$$\frac{1 - \Sigma_i \tau_i \dfrac{\partial y_i}{\partial G} - \Sigma_k t_k \mathrm{E}\left(\dfrac{\partial x_k}{\partial G}\right)}{\mathrm{E}(\alpha)} \frac{\mathrm{d}V^*}{\mathrm{d}\tau_j},$$

with

$$\frac{\mathrm{d}V^*}{\mathrm{d}\tau_j} = \frac{\partial V^*}{\partial \tau_j} + \frac{\partial V^*}{\partial G}\frac{\partial G}{\partial \tau_j}.$$

To see this, totally differentiate (16) with respect to τ_j and simplify to get an expression for $\partial G/\partial \tau_j$; use this and the expressions for $\partial V^*/\partial G$ and $\partial V^*/\partial \tau_j$, from (10a)–(10b), to derive an expression for $\mathrm{d}V^*/\mathrm{d}\tau_j$, then simplify while making use of (14a)–(14b). Note also that $\Sigma_i \tau_i(\partial y_i/\partial G) + \Sigma_k t_k \mathrm{E}(\partial x_k/\partial G)$ is, as usual, the income effect due to the change in the excess burden of the existing taxes in the system.

[14] In fact, it appears that because of the presence of uncertainty and lack of insurance, all the substitution terms on the left-hand sides of (18a)–(18b) have ambiguous signs.

insurance benefit term $-\mathrm{cov}(\alpha, x_s)$ will be positive if x_s is a normal good. In this case, the individual consumes more x_s and pays more taxes in a good state of nature than in a bad one. We have the opposite result if the good is inferior; but it is clear that at least one of the x_ss must be a normal good.

While insurance explains the covariance term that appears in (18b), commitment explains why there is no covariance term in (18a). Everyone spends the same amount of resources on y_js. Thus taxing y_js cannot provide any insurance per se as it affects everyone in the same way.

To further understand the roles that insurance and commitment play in the make-up of optimal taxes consider the following special cases. Assume there is no uncertainty so that there is no need for insurance. In that case $\mathrm{cov}(\alpha, x_s) = 0$ for all s so that $\tau_i = t_k = 0$ is the only solution to (18a)–(18b).[15] This is not surprising. It is well known that without uncertainty, and with identical individuals, the optimal tax policy is to set the tax rates equal to zero and to collect the entire tax revenue through the lump-sum tax G.

With uncertainty, on the other hand, $\tau_i = t_k = 0$ is no longer a solution to the first-order conditions. Distortionary taxes may thus be superior to a lump-sum tax because they provide some insurance to the individual. Note that the need for insurance arises regardless of commitment. Assume the individual need not commit to consumption of any goods. Eqs. (18a)–(18b) will then be symmetric with a covariance term appearing everywhere on the right-hand side. It is again obvious that zero tax rates are no longer a solution to the first-order conditions. However, there will then exist certain restrictions on preferences[16] that call for optimal tax rates to be uniform.

Commitment introduces a fundamental asymmetry between Eqs. (18a) and (18b). It will become apparent in the next section that this asymmetry calls for the optimal tax rates to be non-uniform (even with the restrictions on preferences discussed in the literature). Intuitively, this is due to the insurance that taxation of post-uncertainty (but not pre-committed) goods provide. With commitment, uniform commodity taxation continues to dominate lump-sum taxation; it provides insurance by shifting aggregate expenditures from good to bad states of nature. However, this is *not* optimal when consumption of some commodities is determined before and some after the resolution of uncertainty; the former goods require no insurance. Differential commodity taxes become useful in that they allow an individual to insure himself differently for committed and non-committed goods.

[15] The same is true if preferences are linear in x_ss so that individuals are risk-neural. In that case, α would be independent of w and all covariance terms would be equal to zero.

[16] These are the restrictions discussed by Atkinson and Stiglitz (1976), Atkinson (1977), Deaton (1979) and Deaton and Stern (1986).

4. The usefulness of commodity taxes

The initial doubt about the usefulness of commodity taxes was cast by Atkinson and Stiglitz (1976) who showed that in a number of special cases, differential commodity taxation is unnecessary given an optimal linear income tax. Subsequently, Atkinson (1977) showed that the result is also true with Stone–Geary preferences over leisure and goods; while Deaton (1979) generalized it to the case where preferences are weakly separable between leisure and goods, and where the subutility for goods is such as to ensure linear Engel curves. These results are based on a community with identical tastes but different skills. More recently, Deaton and Stern (1986) generalize the result further by allowing a bit of taste differentiation on the part of the individuals (the case of linear and parallel Engel curves).

This section re-examines the issue of the usefulness of commodity taxes. We prove that with uncertainty, optimal tax rates are non-uniform, thus demonstrating that differential commodity taxation enhances welfare over and above an optimal linear income tax. We establish this result by showing that *if* the optimal tax rates within each category of goods are the same, they must differ between categories. This proves that we can *never* have equal optimal tax rates on all goods. The proof of this, and some later claims, will be made easier if we first establish the following two lemmas.

Lemma 1. (i) A change in q_j must necessarily affect the compensated supply of labor. That is,

$$\frac{\partial L^c}{\partial q_j} \neq 0 . \tag{19}$$

(ii) We have

$$\sum_i q_i \frac{\partial y_i^c}{\partial q_j} \Big/ \frac{\partial L^c}{\partial q_j} > 0 . \tag{20}$$

The proof is given in the Appendix.

Lemma 2. (i) The following relationship holds for all *values of $\underset{\sim}{q}$ and $\underset{\sim}{p}$:*

$$\sum_i q_i \frac{\partial y_i^c}{\partial q_j} + \sum_k p_k \,\mathrm{E}\left(\frac{\partial x_k^c}{\partial q_j}\right) = \frac{\partial L^c}{\partial q_j} \mathrm{E}(w) . \tag{21}$$

(ii) The following relationship holds at the optimal *values of the tax rates:*

$$\sum_i \frac{\partial y_i^c}{\partial q_j} + \sum_k \mathrm{E}\left(\frac{\partial x_k^c}{\partial q_j}\right) = \frac{\partial L^c}{\partial q_j} \mathrm{E}(w) . \tag{22}$$

Proof. To obtain (21) it is sufficient to take the expected value of the budget constraint (2) then differentiate it with respect to q_j and make use of (13a). To derive (22), simply substitute (18a) into (21) while recalling that $q_i = 1 + \tau_i$ and $p_k = 1 + t_k$. □

We are now in a position to state and prove the following proposition:

Proposition 2. The optimal tax rates are non-uniform.

Proof. A necessary condition for the optimal tax rates to be equal is that $\tau_i = \tau$ and $t_k = t$, for all i and k. Assume first that $t = 0$. Clearly in that case, $\tau_i = 0$ is not a solution to Eqs. (18a)–(18b). Now assume $t \neq 0$. Multiplying (22) by $1 + t$ and subtracting (21) from the resulting expression yields

$$(t - \tau) \sum_i \frac{\partial y_i^c}{\partial q_j} = t \frac{\partial L^c}{\partial q_j} \mathrm{E}(w) , \tag{23}$$

which must be satisfied at the optimal values of the tax rates. Since $t \neq 0$, it immediately follows from lemma 1 that $\tau \neq t$. Hence the optimal tax rates cannot be equal. □

Proposition 2 proves that optimal taxation requires a mix of differential commodity taxes and a uniform lump-sum tax. It is important to point out that we have not imposed any additional restrictions on preferences to prove this proposition. *The result thus holds for the class of utility functions characterized by* (1) *which includes all types of preferences for which uniform taxation is optimal in the absence of uncertainty.* Commitment in the presence of uncertainty thus reclaims the lost role of commodity taxes as instruments of optimal tax policy.

The conclusion that restrictions on preferences cannot make commodity taxes redundant may best be understood by considering what such restrictions achieve in the traditional certainty models. Consider, for example, Cobb–Douglas preferences. Atkinson and Stiglitz (1980) have shown, in the context of a certainty model where individuals have identical tastes but differ in earning abilities, that with these preferences optimal commodity taxes are uniform. They explain their result this way: "Where individuals consume goods in the same proportion, it is not possible to use indirect taxes to redistribute – they impose the same percentage burden on everyone" (1980, p. 389). In our setting, however, restrictions on preferences cannot do the trick. Given commitment, individuals will not consume goods in the same proportion: While the expenditures on x_ks are a constant fraction of ex-post income irrespective of the actual realization of the wage, the expenditures on y_is are not. They are determined beforehand and are thus independent of realized income. It is this property that drives our result and reclaims for commodity taxes their lost role.

The foregoing discussion leads one naturally to ask if any concrete results may be obtained concerning the magnitudes of the tax rates that apply to the two categories. Our earlier insurance argument suggests intuitively that x_ks (goods whose consumption levels become known *after* the resolution of uncertainty and thus are, beforehand, random) should be taxed absolutely as well as relative to y_is (goods whose consumption levels are committed to *before* the resolution of uncertainty and as such serve no obvious insurance purpose). However, at this level of generality, it would be impossible to obtain such categorical results. With m goods in one category and n in the other, there will be many different types of interaction between them. It is quite likely that these interactions will mask, and even overshadow, the insurance property that we have been discussing and trying to isolate.

What is needed, in order to concentrate on the interaction between the two categories, is to abstract away from the interactions within the categories. This is achieved most easily by assuming that there is only one good in each category ($m = n = 1$). With more than one good per category, a meaningful comparison appears to be possible only if there is one tax rate for each category. This may be achieved in two ways. First, one can restrict preferences in such a way that optimal tax policy implies that goods within each category should be taxed uniformly. For example, one can prove that when preferences are Cobb–Douglas (so that commodity taxation is unnecessary in the traditional model), optimal taxation calls for goods within each category to be taxed uniformly. (See Cremer and Gahvari, 1994). Second, instead of restricting preferences, one can ensure uniform taxation within categories by restricting the type of taxes that are available to the government.

It will become clear below that either avenue will yield the same set of results. However, the exposition will be made simpler if we begin our discussion by ignoring the last possibility. We can then come back to it at the end of the section. We are now in a position to prove two further propositions.

Proposition 3. If $n = m = 1$, or if the optimal tax policy implies that x_ks and y_is are taxed uniformly at rates t and τ, then $t \geq 0$.

Proof. We show that if a tax policy involves $t < 0$ it can always be dominated by an alternative tax policy. Denote the variables under the *initial* policy by an overbar and assume $\bar{t} < 0$. Replace the initial tax rates by $t = \tau = 0$ and adjust the lump-sum term according to

$$G = \bar{G} - \bar{t} \sum_k \mathrm{E}(\bar{x}_k) - \bar{\tau} \sum_i \bar{y}_i \tag{24}$$

so that the expected tax revenue remains unchanged regardless of the individual's actual choice under the new policy. Now because both policies yield the same expected tax revenue, the new policy will be better *if it allows the consumer to achieve a higher level of expected utility*. We show that this is the case *even if he continues to choose* $\bar{y}_i$ $(i = 1, 2, \ldots, m)$ and $\bar{L}$.[17]

Under the above assumption, it is easily seen from (3), (24), and $t = \tau = 0$ that I is related to $\bar{I}$ via

$$I(w) = \bar{I}(w) - \bar{t} \sum_k \mathrm{E}(\bar{x}_k) . \tag{25}$$

Moreover, under the initial policy, we have that $\bar{I} = (1 + \bar{t}) \Sigma_k \bar{x}_k$. Taking the expectation of this expression and substituting into (25), we can rewrite it as

$$I(w) = \bar{I}(w) - \frac{\bar{t}}{1 + \bar{t}} \mathrm{E}(\bar{I}(w)) . \tag{26}$$

Next, given that the individual is choosing $\bar{L}$, $\bar{y}_1$, $\bar{y}_2$ and $\bar{y}_m$ under both policies, it follows from (7) that the change in his expected utility can be calculated by comparing $\mathrm{E}[v(1 + \bar{t}, 1 + \bar{t}, \ldots, 1 + \bar{t}, \bar{I})]$ under the old tax regime to $\mathrm{E}[v(1, 1, \ldots, 1, I)]$ under the new regime. But, the zero homogeneity of v implies that

$$\mathrm{E}[v(1 + \bar{t}, 1 + \bar{t}, \ldots, 1 + \bar{t}, \bar{I})] = \mathrm{E}\left[v\left(1, 1, \ldots, 1, \frac{\bar{I}}{1 + \bar{t}}\right)\right] .$$

This together with (26), relating I linearly to $\bar{I}/(1 + \bar{t})$ and indicating that $I(w)$ has the same mean as $\bar{I}(w)/(1 + \bar{t})$ but a lower variance, and the fact that the individual is risk averse implies that the expected utility under the new tax regime has increased. □

This result is intuitive and very interesting. As has been argued before, the reason for imposing commmodity taxes in our model is to provide insurance. As insurance is provided through x_ks only, one intuitively expects these goods to be taxed and not subsidized.[18]

The insurance argument also suggests that x_ks should be taxed at a higher rate than y_is. This follows because the sole reason for taxing y_is is to counteract the substitution effects of the tax on x_ks. Proposition (4) confirms this intuition.

[17] This is sufficient for our purpose. Of course, he can do even better under the new policy if these variables are also adjusted.

[18] We will show below that t must in fact be *strictly* positive.

Proposition 4. If $n = m = 1$, or if the optimal tax policy implies that x_ks and y_is are taxed uniformly at rates t and τ, then $t > \tau$.

Proof. From part (ii) of lemma 1 we have

$$\sum_i q_i \frac{\partial y_i^c}{\partial q_j} \bigg/ \frac{\partial L^c}{\partial q_j} > 0 \, .$$

Replacing q_i with $1 + \tau$, substituting from (23) and simplifying yields

$$\frac{(1+\tau)t}{t-\tau} \mathrm{E}(w) > 0 \, . \tag{27}$$

It then follows from (27) and proposition 3 that $t > \tau$. □

We can also use proposition 4 to show that t is in fact strictly positive, and not just non-negative as proved in proposition 3. If $t = 0$ it follows from proposition (4) that $\tau < 0$. This in turn implies that $G < 0$. However, this cannot happen, given risk aversion and the need for insurance. The result is summarized in the following corollary.

Corollary. If $n = m = 1$, or if the optimal tax policy implies that x_ks and y_is are taxed uniformly at rates t and τ, then $t > 0$.

The implications of proposition 4 and the above corollary are far reaching and most interesting. Recall our earlier discussion of the nature of the commodities that constitute our two categories. As was discussed then, goods whose consumption levels are committed to before the resolution of uncertainty (y_is) are goods such as housing purchases and consumer durables; while the other category (x_ks) consists of non-housing and non-durable purchases. Our findings thus have the interesting implication that the purchases of housing and other durables should be subsidized relative to non-housing and non-durable goods. This is a most interesting implication in light of all the literature criticizing the implicit housing subsidies that are embedded in the US and the UK tax systems (see Rosen, 1985).

Finally, we observed in our discussion preceding proposition (3) that one way to ensure that there is a single tax rate per category is to restrict the set of tax instruments. We are now in a position to discuss the implications of this possibility. Consider again the case where m and n are >1 and preferences are represented by (1). Assume further that the tax instruments available to the government consist of a single tax rate per category. Setting up the government's problem as in Section 3, one can easily show that the optimal (restricted) tax rates will now be characterized by a set of two equations similar to (18a)–(18b) with τ_is being replaced by τ and t_ks by t. It will then also be simple to see that the proofs of propositions 3 and 4, and the corollary continue to go through. This is summarized as proposition 5.

Proposition 5. If the tax rates on x_ks and y_is are restricted *to be uniform at rates t and τ, then optimal taxation calls for $t > 0$ and $t > \tau$.*

5. Concluding remarks

The paper has re-examined the theory of optimal commodity taxation in a framework which explicitly allows for uncertainty. Two important lessons have emerged. The first is that differential commodity taxes *do* have a role to play as instruments of optimal tax policy – an optimal linear income tax will not suffice. The result is quite robust; it does not go away by imposing restrictions on preferences as long as consumers are risk averse. This is quite significant, particularly in the light of recent attempts to demonstrate that commodity taxes are, under some circumstances, unnecessary in the presence of an optimal linear income tax.

The second lesson is that goods with consumption levels which are determined prior to the resolution of uncertainty (housing purchases) should face a tax rate which is lower than the tax rate on goods with consumption levels determined after (purchases of non-housing goods). This provides an argument in favor of housing subsidies.

The lessons have been drawn on the basis of several assumptions: wage uncertainty, commitment, linear commodity taxes and a uniform lump–sum tax as the only available tax instruments, and the separability of preferences. The paper has throughout emphasized the importance of uncertainty and commitment. Uncertainty calls for taxation to provide insurance while commitment implies that optimal tax rates are non-uniform (with pre-committed goods taxed at a lower rate). We may now briefly examine the significance of the other assumptions to our results.

The paper has assumed that in addition to consumption of some goods, individuals also commit to labor supply. However, our results apply equally to the cases where the uncertainty environment modelled does not call for labor supply to be pre-committed. All the results of the paper remain valid in both situations.[19]

[19] To show this, simply follow the steps of Sections 2 and 3 of the paper and observe that dropping the pre-commitment assumption will change only the specification of the individual's first-order condition for the labor supply [Eq. (8a)]. Note in particular that the specification of the government's problem and its solution remain intact. We have assumed pre-commitment to labor for conformity with Eaton and Rosen (1980a). Persson (1983), on the other hand, assumes no pre-commitment. He also points out the equivalence of the two approaches.

The observation that pre-commitment to labor supply is not important, while pre-commitment to goods is, should not be surprising. The point is that if an individual commits to consumption of a particular good, he will be spending a fixed amount of money on it regardless of his future realization of wage. Hence taxing such a good raises the same amount of revenue from everyone and will serve no insurance purpose. This is not the case with taxation of pre-committed labor. With the wage being uncertain, a wage tax serves as a source of insurance whether or not one pre-commits to labor supply.[20]

As may be expected, our results depend on the availability of tax instruments. We have ruled out a general income tax including taxation contingent on w. If labor supply is pre-committed, this latter type of taxation may in fact be feasible. In this case, any differences in individuals' realized incomes must be due to the wage realization as everybody chooses the same level of labor supply. It follows that if incomes are observable, realized wages must also be observable. Taxation contingent on w will thus be lump-sum and first-best efficient. (This possibility has been discussed in Lundholm, 1992).

The results of the paper also depend on the separability of preferences between committed and non-committed goods. Without separability, the characterization of the optimal tax rates will be more complicated. However, if it were not for separability, the question of the ineffectiveness of commodity taxes would not even arise. Note that it is the separability between committed (including labor supply) and non-committed goods that is important here. The separability of labor supply and the rest of the pre-committed goods is not crucial.[21]

Finally, it must be pointed out that the paper has not addressed two particularly important issues. First, we have left open the question of the usefulness of commodity taxes in the presence of a general income tax. We plan to examine this issue in a sequel to this paper. Second, we have not discussed the question of randomness in prices. This is particularly important for pre-committed goods such as housing. The type of uncertainty we have modelled varies among individuals, thus ruling out randomness in prices. This is a very interesting question requiring a different modelling strategy; it is left for future research. The direct versus indirect tax controversy is far from settled.

[20] It must also be pointed out that if labor supply were supplied exogenously, the tax system can be used to provide full insurance without creating any welfare costs. As in Eaton and Rosen (1980a), this may be done by levying a 100% wage tax coupled with lump-sum cash payments. Obviously, with full insurance, differential commodity taxation will not be useful.

[21] In fact, we use this particular form of separability – as stipulated in Eq. (1) – in the paper only in the proof of lemma (1). Given the lemma, *all* the results of the paper hold for preferences that are separable only between committed and non-committed goods.

Acknowledgements

We thank Jacques Crémer, Jim Poterba, the seminar participants at the Institut für Finanzwissenschaft und Steuerrecht, Vienna, Austria, and particularly two referees for helpful comments.

Appendix

Proof of lemma 1.
(i) The proof is by contradiction. First, substitute L^c and y_i^c for L and y_i in the first-order conditions (8a)–(8b), differentiate with respect to q_j ($j = 1, 2, \ldots, m$) and simplify to get

$$\left[\mathrm{E}\left(w \frac{\partial \alpha}{\partial I}\right)\right] \sum_i q_i \frac{\partial y_i^c}{\partial q_j} = \left[\varphi'' + \mathrm{E}\left(w^2 \frac{\partial \alpha}{\partial I}\right)\right] \frac{\partial L^c}{\partial q_j}, \tag{A1}$$

$$\sum_i \left[\psi_{li} + q_l \,\mathrm{E}\left(\frac{\partial \alpha}{\partial I}\right) q_i\right] \frac{\partial y_i^c}{\partial q_j} = q_l \,\mathrm{E}\left(w \frac{\partial \alpha}{\partial I}\right) \frac{\partial L^c}{\partial q_j},$$
$$l = 1, \ldots, j-1, j+1, \ldots, m, \tag{A2}$$

$$\sum_i \left[\psi_{ji} + q_j \,\mathrm{E}\left(\frac{\partial \alpha}{\partial I}\right) q_i\right] \frac{\partial y_i^c}{\partial q_j} = q_j \,\mathrm{E}\left(w \frac{\partial \alpha}{\partial I}\right) \frac{\partial L^c}{\partial q_j} + \mathrm{E}(\alpha). \tag{A3}$$

Now suppose $\partial L^c / \partial q_j = 0$. Substituting into (A1)–(A3) and simplyfing yields

$$\sum_i q_i \frac{\partial y_i^c}{\partial q_j} = 0, \tag{A4}$$

$$\sum_i \psi_{li} \frac{\partial y_i^c}{\partial q_j} = 0, \quad l = 1, \ldots, j-1, j+1, \ldots, m, \tag{A5}$$

$$\sum_i \psi_{ji} \frac{\partial y_i^c}{\partial q_j} = \mathrm{E}(\alpha). \tag{A6}$$

Assuming equations (A4)–(A5) are linearly independent,[22] they imply

$$\frac{\partial y_1^c}{\partial q_j} = \frac{\partial y_2^c}{\partial q_j} = \cdots = \frac{\partial y_m^c}{\partial q_j} = 0,$$

so that from (A6), $\mathrm{E}(\alpha) = 0$, which is not the case.

[22] Linear independence is needed only for *one* j. If this is not satisfied, the Hessian matrix associated with ψ must always be singular.

(ii) This result is immediate from (A1) and the facts that $\varphi'' < 0$ and $\partial\alpha/\partial I < 0$ (from concavity of u).

References

Atkinson, A.B., 1977 Optimal taxation and the direct versus indirect tax controversy, Canadian Journal of Economics, 10, 590–606.

Atkinson, A.B. and J.E. Stiglitz, 1976 The design of tax structure: direct versus indirect taxation, Journal of Public Economics, 6, 55–75.

Atkinson, A.B. and J.E. Stiglitz, 1980 Lectures on Public Economics. New York: McGraw-Hill.

Cremer, H. and F. Gahvari, 1994 Uncertainty and optimally uniform commodity taxes, mimeo.

Deaton, A., 1979 Optimally uniform commodity taxes, Economics Letters, 2, 357–361.

Deaton, A. and N. Stern, 1986 Optimally uniform commodity taxes, taste differences and lump-sum grants, Economics Letters, 20, 263–266.

Diamond, P.A., L.J. Helms and J.A. Mirrlees, 1980 "Optimal taxation in a stochastic economy: A Cobb–Douglas example," Journal of Public Economics, 14, 1–29.

Eaton, J. and H.S. Rosen, 1980a Labor supply, uncertainty, and efficient taxation, Journal of Public Economics, 14, 365–374.

Eaton, J. and H.S. Rosen, 1980b Optimal redistributive taxation and uncertainty, Quaterly Journal of Economics, 94, 357–364.

Eaton, J. and H.S. Rosen, 1980c Taxation, human capital, and uncertainty, American Economic Review, 70, 705–715.

Hamilton, J., 1987 Opitmal wage and income taxation with wage uncertainty, International Economic Review, 28, 373–388.

Kaplow, L., 1991 A note on taxation as social insurance for uncertain labor income, NBER Working Paper 3708.

Koskela, E., 1987 Changes in tax progression and labour supply under wage rate uncertainty, Public Finance/Finances Publiques, 42, 214–226.

Lundholm, M, 1992 Efficient taxation under wage rate uncertainty, Public Finance/Finaces Publiques, 47, 248–256.

Mazur, J., 1989 Optimal linear taxation with stochastic incomes, Public Finance/Finances Publiques, 44, 31–50.

Mirrlees, J.A., 1990 Taxing uncertain incomes, Oxford Economic Papers, 42, 34–45.

Persson, M., 1983 The distribution of abilities and the progressive income tax, Journal of Public Economics, 22, 73–88.

Richter, W.F., 1987 Taxation as insurance and the case of rate differentiation according to consanguinity under inheritance taxation, Journal of Public Economics, 33, 363–376.

Rosen, H.S., 1985 Housing subisidies: effects on housing decisions, efficiency, and equity, Handbook of Public Economics Vol. 1, ed. Auerbach, A. and M. Feldstein, Amsterdam, North-Holland, 375–420.

Stiglitz, J.E. 1975 Incentives, risk, and information: notes towards a theory of hierarchy, Bell Journal of Economics, 6, 552–579.

Varian, H.R., 1980 Redistributive taxation as social insurance, Journal of Public Economics, 14, 49–68.