Public Finance Research Papers (ISSN 2284-2527)

Dipartimento di Studi Giuridici
Filosofici ed Economici

Sapienza
Università di Roma

Istituto di Economia e Finanza

Public Finance Research Papers

A Computational Model of Optimal Commodity Taxation

John T. Revesz

E-PFRP N. 4

2014

John T. Revesz
Former Research Economist, Australian Public Service
Email: john.revesz2@upcmail.hu

Si prega di citare così: John T. Revesz (2014), "A computational model of optimal commodity taxation", Public Finance Research Papers, Istituto di Economia e Finanza, DIGEF, Sapienza University of Rome, n. 4 (http://www.digef.uniroma1.it/ricerca).

John T. Revesz

A computational model of optimal commodity taxation

Abstract

This report examines the structure of optimal commodity tax rates in a many-person many-goods static computational model using segmented LES utility. One of the major findings is that with non-linear Engel curves and linear income tax, optimal commodity tax rates tend to be progressive and highly dispersed under logarithmic utility specifications. However, the dispersion of tax rates is considerably reduced if the inequality aversion of society is low or if tax evasion depends among other things on disparities between commodity tax rates. With exogenously given non-optimal and non-linear income tax schedules, usually there is still a need for differentiated and progressive commodity taxation. Tax evasion tends to reduce optimal tax rates for necessities but increases them for luxuries. Private compliance costs and government administration costs reduce optimal tax rates by a similar amount to the share of these costs from taxes. The results indicate that in a redistributive model the effect of externalities on optimal tax rates exceeds the corresponding Pigovian tax rates or subsidies. The main benefit of higher taxes on leisure complements than leisure substitutes appears to relate to increased tax revenue for redistribution rather than improvement in the utility position of those paying the taxes. The effect of complexities such as tax evasion, administrative costs, externalities and leisure complements/substitutes on redistribution is not neutral. Generally, these complexities tend to increase the progressivity of optimal commodity tax rates. Explanations are provided why the numerical results presented here do not contradict the Laroque-Kaplow proposition, which advocates uniform commodity taxation. Some practical application problems and logical weaknesses of the Laroque-Kaplow proposition are noted.

JEL Classification: C63, H21.

Keywords: optimal taxation; computational models.

1. Outline of the report

The purpose of this paper is to examine the pattern of optimal commodity tax rates using a computational model. The discussion covers among other things the impact on optimal tax rates of administrative and compliance costs, tax evasion, externalities and leisure substitution or complementarity.

The numerical results from this static model suggest that given non-linear Engel curves and logarithmic utility, commodity tax rates tend to be progressive and widely dispersed. Adding “real life” complexities to the model increases further the progressivity and dispersion of tax rates. These findings are in marked contrast to the continuing preoccupation of much of the optimal taxation literature with uniform commodity taxation, which in my view applies only under highly abstract and unrealistic conditions. The report is structured as follows:

Section 2 reviews briefly the history of the controversy on whether commodity tax rates should be uniform or otherwise, and the current state-of-affairs in this debate. This is followed by a review of previous computational studies on optimal commodity taxation. The review suggests that this numerical study is more comprehensive than anything that was done on this subject before.

Section 3 describes the main mathematical features of the segmented Linear Expenditure System (LES) model. Segmented LES means two LES functions with different parameters for two population groups. This arrangement is needed in order to obtain non-linear Engel curves with LES. Attention is given to the fact that while many functional forms could be used to investigate optimal commodity tax rates, segmented LES, which has explicit demand formulas, presents probably one of the easiest options to attack the problem. For the purpose of analysing the numerical results, we develop in appendix 2 an approximate formula for optimal tax rates called the modified inverse elasticity rule. It is used to explain and analyse the numerical results. It is not applied to calculate optimal tax rates, which are found using iterative calculations.

Section 4 examines the role of commodity taxation in income redistribution. The model can incorporate various inequality aversion rates to represent political value judgments concerning income distribution. The numerical results indicate that given non-linear Engel curves and linear income tax, generally luxuries should be taxed at much higher rates than necessities. However, if the inequality aversion is low then the dispersion of optimal tax rates is considerably reduced. We also examine the structure of optimal commodity tax rates associated with some exogenously given non-linear and non-optimal income tax schedules.

Section 5 deals with tax evasion and administration. The modified inverse elasticity rule is extended to accommodate these factors. The results reveal certain dichotomy between necessities and luxuries, As a result of tax evasion optimal tax rates on necessities are reduced or increased only slightly, while those on luxuries are significantly increased. We also examine the possibility that evasion may depend by how much a commodity’s tax rate exceeds the average tax rate. Introducing a tax rate dependent evasion factor can cause a large reduction in the dispersion of optimal tax rates. Public administration and private compliance costs cause a reduction in optimal tax rates by a similar amount to the share of these costs from commodity taxes.

Section 6 examines the impact of negative or positive externalities on the solution. The general conclusion is that in a redistributive model their impact on optimal tax rates or subsidies exceeds the respective Pigovian tax rates or subsidies. The analysis in this section follows a similar line to Sandmo (1975).

Section 7 looks at the impact on optimal tax rates of paternalistic concerns. These include a number of taxes and subsidies provided in line with political judgments about what is in the interest of the consumer in the long run. Examples include “sin goods”, home buying, educational books and software or expenditure on preventative health care.

Section 8 examines the impact of leisure substitution or complementarity on optimal tax rates. We focus for that purpose on the cross-derivatives of labour supply with respect to the price of goods. Using the modified inverse elasticity rule, an approximate formula is worked out for the impact of leisure non-separability on optimal tax rates.

Section 9 examines the proposition put forward by Laroque (2005) and Kaplow (2006) that given weakly separable utility and identical preferences, non-uniform commodity taxes could be replaced in a Pareto improving manner by uniform commodity taxes combined with compensating adjustment to the income tax schedule. Explanations are given why our numerical results do not contradict mathematically the Laroque-Kaplow (LK) proposition. Some limitations of the LK proposition will be also examined.

Section 10 presents summary and qualifications.

Appendix 6 illustrates the numerical impact on optimal tax rates of some other factors incorporated into the model. These include expenditure on public goods and targeted support to the needy. The consideration given to these factors is more limited than in the text.

The source code for this computational model is located in the website: http://john1revesz.com It can be run and explored by the interested reader using a QBasic compiler that is available on the Internet free of charge. A short user’s guide is presented in Appendix 1. A more extensive guide is located in the website. The computer model can incorporate a wide range of specifications in regard to utility parameters, wage distribution, the inequality aversion of society, tax evasion, administrative and compliance costs, externalities, non-linear income tax, public goods expenditure requirements, substitution or complementarity with leisure and other factors that will be described later.

2. Historical background

The theory of optimal taxation can be dated back to Ramsey’s work in the 1920’s, but only since the early 1970s were mathematical models developed that focus on the redistributive aspects of taxation. This theory considers both equity and efficiency. In these models individuals differ in term of labour productivity (ability) and redistributive taxation is used in line with the inequality aversion of society. The pioneering model in this field is Mirrlees (1971), who examined how to optimise a non-linear income tax function combined with a uniform lump-sum grant (called the demogrant), taking into account that more income redistribution will cause a reduction in labour supply.

Some of the early studies on commodity taxation in redistributive models favoured the idea of progressive indirect taxation. Feldstein (1972) showed that the pricing of goods supplied by public agencies should be progressive. Diamond (1975) found that in the presence of a redistributive demogrant, commodity tax rates will tend to be progressive. Neither of these studies included explicitly labour supply.

Shortly after the publication of Diamond’s article, a growing emphasis emerged in the literature on conditions that lead to uniform optimal commodity tax rates. These conditions include weakly separable utility between commodities and leisure, perfect competition, no differentiation in support payments, no administrative-compliance costs, no tax evasion and consumers with identical characteristics apart from the wage rate. Atkinson and Stiglitz (1976) extended the Mirrlees (1971) income tax model and showed that provided the income tax is the optimal control theoretic solution then optimal commodity tax rates will be uniform. Christiansen (1984) extended the Atkinson-Stiglitz result to a simpler model where commodity tax rates are proportional. Deaton (1979) showed that with linear Engel curves and linear income tax, optimal commodity tax rates will be uniform. More recently, Laroque (2005) and Kaplow (2006) suggested that optimal commodity tax rates should be uniform, even if the income tax function is not optimal to start with. The only significant qualification to uniformity presented in these articles is that utility should be weakly separable between commodities and leisure. Otherwise, leisure-substitutes should be taxed more lightly and leisure complements more heavily.

At an early stage, the main qualification to the uniform tax proposition was concerned with differentiated lump-sum support grants. Deaton and Stern (1986) showed that if differentiated support grants are set at a sub-optimal level then non-uniform commodity taxes may be justified. Their proposition was tested and confirmed by Ebrahimi and Heady (1987) through a computational study, using child benefits as the reason for differentiating lump-sum grants.

Another objection to commodity tax uniformity is concerned with the infeasibility of providing optimal income subsidies in certain situations. This is particularly relevant in many developing countries where direct support payments are absent. In this situation only progressive indirect taxation can be used to provide some support to the needy, as limited as it may be. Computational studies by Ray (1986), Murty and Ray (1987) and Srinivasan (1989), show that given inequality aversion, in the absence of direct support the optimal commodity tax structure will be progressive, with higher taxes on luxuries and lower taxes or subsidies on necessities.

In the 1990s the controversy about uniform commodity taxation has developed further. A number of mathematical studies pointed out the restricted validity of the uniform tax proposition, besides leisure substitution or complementarity. Boadway, Marchand and Pestieau (1994) looked at the effect of tax evasion and concluded that in the presence of income tax evasion optimal commodity tax rates will not be uniform, unless commodity preferences are quasi-homothetic. Later other factors were noted that may invalidate the uniform tax proposition. These include Cremer and Gahvari (1995) on uncertainty and home purchases, Myles (1995) on imperfect competition, Naito (1999) on non-linear technologies and imperfect mobility between skills, Cremer, Pestieau and Rochet (2001) on differences in earning characteristics other than the wage rate (such as wealth) and Saez (2002) on heterogeneous preferences between households. In a critical review of earlier studies, Alm (1996) questioned the practical relevance of the uniform tax propositions derived from highly simplistic models. He noted that the administrative and compliance costs of taxes are often comparable or even larger than the economic efficiency costs measured in terms of excess burden. Boadway and Pestieau (2003) pointed out that the Atkinson-Stiglitz (1976) uniformity theorem is only valid if the income tax function is the Mirrlees (1971) type optimal solution. Thus given actual income tax schedules, the conditions for commodity tax uniformity may not apply.[1] They also examined other sources of violations due to different needs and endowments, multiple forms of labour supply and home production.[2] A recent publication by Bastani, Blomquist and Pirttila

[1] The Laroque-Kaplow proposition (discussed in section 9) suggests that the Atkinson-Stiglitz result can be extended to non-optimal income tax schedules.

[2] I have also published some papers criticising the uniform tax proposition (see Revesz (1986, 1997, 2005)). The main objections raised in these papers will reappear in a different form here.

(2013) examines the need for child care subsidies and finds that commodity tax rates should be progressive. [3]

Despite numerous objections and qualifications, the uniform commodity tax proposition seems to be alive and well. The Mirrlees et. al. (2011) review of the UK tax system concluded that apart from child care and externalities there is no reason for sustaining the existing differentiation in the VAT system, provided that any moves toward uniformity are compensated through appropriate adjustments to income tax.

The continuing preoccupation with commodity tax uniformity may partly explain the paucity of computational research on optimal commodity taxation. A literature review on this subject is presented by Nygard (2008). Most of the computational studies on optimal commodity taxation deal with the single-person Ramsey model and are largely irrelevant to redistributive taxation.

The few studies that have been published on many-person models, deal mainly with the situation in developing countries where direct support payments are absent (see Ray (1986), Murty and Ray (1987) and Srinivasan (1989)). In the presence of egalitarian objectives and in the absence of other means of support (such as a demogrant), progressive commodity taxation can always be justified, regardless of whether or not labour supply is present in the model.[4] As far as I can see, apart from Murty and Ray (1987), labour supply does not play a role in these models. While these studies are important in their own right, they bear little relevance to the tax uniformity debate in developed countries, where targeted grants and other means of direct support are widely employed.

Ebrahimi and Heady (1987) paper on the influence of demographic variables on optimal commodity tax rates is perhaps the most sophisticated and comprehensive computational model among early studies. Ebrahimi and Heady (1987) examine the effect of providing differentiated support payments depending on the number and age of children in the household. They also examine differences in male and female labour supply. They find that non-uniform commodity tax rates are justified when:

I. commodities and leisure are not weakly separable,
II. Engel curves are not parallel across households,
III. the demogrant per household is not linked to the age and number of children.

Their model is based on econometric estimates. It is restricted to only four commodities (energy, food, clothing and other goods), as well as female and male leisure. They did not examine non-linear Engel curves, but examined non-parallel Engel curves across households. The heterogeneity of Engel curves causes some departure from strict linearity. Even that slight non-linearity of Engel curves caused a perceptible differentiation in optimal tax rates. They also found that if child benefits are set at a sub-optimal level then optimal commodity tax rates will be differentiated.

Revesz (1997) describes a fairly extensive computational model of optimal commodity taxation, but unfortunately that paper has been largely ignored in the literature. Revesz (1997) provides the foundation for the model discussed in the present study.

---

[3] It should be noted that apart from Deaton and Stern (1986), Ebrahimi and Heady (1987), Naito (1999) and Saez (2002), all these mathematical studies rely on the Stiglitz (1982) self-selection model. This is a two-person or a few persons (or groups) approach to tax optimisation that can yield analytical insights. It is better suited for analytical investigations than to computational work in a many-person model.

[4] If variable labour supply is absent from the model then income is fixed. In this situation, in a static model in the presence of a demogrant (b), the optimal policy would be to tax away all incomes (say, by raising commodity tax rates to infinity) and redistributing the tax collected through ‘b’, because despite 100% taxation, national output would be unaffected.

A recent computational study by Bastani, Blomquist and Pirttila (2013) examines the effect of strong leisure substitutes, such as child-care and aged-care services on optimal commodity tax rates. They examine the role of tagging, optimal taxes-subsidies and changes in labour supply, using a Stiglitz (1982) type self-selection model, involving two composite goods plus child care and leisure and four population groups – low and high wage earners, parents and non-parents. They find that provided child care is not fully paid by the government, progressive taxation of commodities is justified.

It should be noted that in contrast to commodity taxation, there is no shortage of computational research on the optimal income tax. The Mirrlees (1971) non-linear income tax model was published together with numerical results, and various extensions of the model have been subject later to further computational studies. The linear income tax model was examined through numerical studies by Stern (1976) and others. Needless to say, the extensive numerical research on optimal income taxation and the paucity of commensurate research on commodity taxation has created a rather unbalanced situation in the theory of optimal taxation.

3. The mathematical framework of the segmented LES model

The basic structure of the computational study discussed in this paper follows the specifications used in the commodity taxation combined with linear income tax models of Deaton (1979) and Atkinson and Stiglitz (1980). The uniform tax solution is used as the benchmark for investigating departure from uniformity under non-linear Engel curves demand conditions, or as a result of the inclusion of other factors into the model.

The government's problem is to maximise the social welfare function:

$$U = \sum_h a_h u_h (\boldsymbol{q_h}) = \sum_h a_h u_h (\boldsymbol{p}, W_h, y_h) \qquad (1)$$

where u_h is the direct or indirect utility function of taxpayer h and a_h represents politically determined utility weights. W_h is the gross wage rate (or ability level) of taxpayer h and y_h is his/her lump-sum income. Maximisation is carried out subject to the revenue constraint

$$\sum_h \sum_i t_i \check{p}_i q_{ih} - Hb - R_0 = 0 \qquad (2a)$$

and the production possibilities constraint

$$\sum_h W_h \ell_h - \sum_h \sum_i \check{p}_i q_{ih} - R_0 = 0 \qquad (2b)$$

where q_i are commodities, $\check{p}_i$ producer prices, t_i commodity tax rates, R_0 is fixed public goods expenditure requirement, H the total number of taxpayers and b is a uniform lump-sum grant per taxpayer (called the demogrant). Producer prices are fixed. Setting $\check{p}_i$ to the numeraire value one, consumer prices are given as: $p_i = 1 + t_i$

At the start the model is confined only to commodity taxation - income taxation will be introduced later. This will be referred to as the baseline model. Demographic issues, such as household composition are ignored in the model. The demogrant is the same for all taxpayers. Further, it is assumed that taxpayers differ in their earning capacity, represented by the gross wage rate W. Assuming all income comes from wages, income (m) is the same as output and is given by $W\ell$, where ℓ is labour supply. Gross income in the post-tax situation after receiving the demogrant (b) is: $\widehat{m} = W\ell + b$. For the time being, lump-sum income 'y' equals 'b'. Various additions to 'y' will be introduced later. Preferences are assumed to be weakly separable, which means that utility u = u(v($\mathbf{q}$), L) , where $\mathbf{q}$ is the vector of commodities, L is leisure and v is a sub-utility function of commodities.

The model employs Linear Expenditure System (LES) utility. The LES utility function is defined as:[5]

$$u = \sum_i \beta_i \log(q_i - \alpha_i) \tag{3}$$

The corresponding demand functions for commodities are:

$$q_i = \alpha_i + \frac{\beta_i}{p_i}\left(Zw + y - \sum_j p_j \alpha_j\right) \tag{4}$$

Labour supply is given as:

$$\ell = Z - q_L = Z(1 - \beta_L) - \alpha_L - \frac{\beta_L}{w}(y - \sum_j p_j \alpha_j) \tag{5}$$

where w is net after income-tax wage rate and y is net after income-tax lump-sum income. Z represents the total time available and q_L is leisure. Given that income tax is not yet present, w = W and y = b.

The only constraint on the parameters is that $\sum_i \beta_i = 1$. The α_i can be positive or negative, but with negative α_i care must be taken to ensure that q_i will not turn out to be negative at low income levels. By virtue of being an additive utility function, LES satisfies weak separability between commodities and leisure. Additivity also implies that LES is globally quasi-concave.[6]

A unique feature of the model is the incorporation of a segmented utility function in order to obtain non-linear Engel curves. The segmentation is defined as follows. There are 15 taxpayers in the model. It is assumed that the eight lower W taxpayers consume only 9 goods (the necessities). The higher W seven taxpayers consume 18 goods, including the 9 necessities plus 9 luxuries. In order to obtain non-linear Engel curves, the β_i parameters of the two groups must be different. In the original model reported in Revesz (1997) the β_i parameters of the second group are obtained by using a so-called splitting factor (denoted s). Denoting the β_is of the eight low W taxpayers as β_{i0}, then the β_is of necessities of higher W taxpayers will be $(1\text{-}s)\beta_{i0}$ and the β_is of luxuries will be $s\beta_{i0}$. This splitting arrangement ensures that the sum of the β_is of the higher W group equals one, and the consumption patterns of the two groups nearly mesh at the border between the eighth and the ninth taxpayer. This splitting arrangement has been retained in the current version of the program, but now the user can also freely specify β_is for the second taxpayer group, even if it does not ensure near meshing of consumption at the border. The splitting arrangement usually leads to similar demand elasticities and optimal tax rates within each group. It will be referred to as the bipolar scenario.[7] The less restrictive arrangement leads to more dispersed demand elasticities and optimal tax rates within each group. It will be referred to as the dispersed parameters scenario.[8] The bipolar and dispersed parameters are defined so as to ensure that the consumption of luxuries starts at the border W between the two groups and there is no negative consumption of any commodity.

Revesz (1997) also reported results from a 9 commodity model where no segmentation was used, and all 15 taxpayers consumed only the 9 necessities and shared

[5] LES appears in many publications. For one, see Thomas (1987). The present version uses instead of income (m) the earning parameters w and y, as well as total time (Z) and leisure (L).

[6] With LES utility constraint (2b) is redundant. The LES demand equations are defined so that the budget equation $\sum_i p_i q_i - (W\ell + b) = 0$ is always satisfied. Summing over h, it is not difficult to see that in this situation provided constraint (2a) is satisfied so will (2b).

[7] All the bi-polar scenarios reported in this paper use a split factor of 0.99. The same applies in Revesz (1997), which used exclusively bi-polar scenarios. Public goods expenditure requirement was not included in Revesz (1997), therefore R_0 was effectively zero.

[8] The interested reader can find the $\propto_i$ and β_i parameters of the bi-polar and dispersed parameter scenarios in the DATA lines of the source program located in the website: http://john1revesz.com Data on the bi-polar scenario was presented in Revesz (1997). The same data are used here.

all utility parameters. The purpose was to test Deaton (1979) theorem that under weakly separable utility, linear income tax and linear Engel curves, optimal commodity tax rates will be uniform. This has been indeed confirmed by the numerical results, though the 9 commodity linear Engel curves preferences turned out to be nearly homothetic, something that is not supported by empirical evidence in a multi-commodity framework. In this paper we shall be concerned mainly with results from the 18 commodity non-linear Engel curves model.

In regard to the choice of LES as the utility function in the model, let me note that I am not aware of any direct utility function apart from LES, CES and the homogeneous Cobb-Douglas function that yields explicit demand formulas. Neither CES nor Cobb-Douglas seems suitable for a multi-product flexible utility model, which leaves LES as the best choice. Needless to explain, the availability of explicit demand formulas makes it much easier to find the numerical optima. Segmented LES is well suited to examine a wide range of non-linear Engel curves configurations. The indirect utility functions employed in earlier computational optimal commodity tax studies involving labour supply (Ebrahimi and Heady (1987), Murty and Ray (1987)), can represent non-linear Engel curves, but at a more limited range than segmented utility.

In the present model the search for the optimum is based exclusively on a computational approach, without any formulas apart from the utility and demand functions. The value of the utility function described in (1) and (3) is evaluated at each step, using the demand functions defined in (4) and (5). At each step (k) of the calculations a new demogrant (b) is found using constraint (2a). Successive approximations to optimal tax rates are carried out using the gradient equation:

$$t_i(k) = t_i(k\text{-}1) + {}^{C^*\Delta U(k)}/_{\Delta t_i{}^*q_i} \qquad (6)$$

where U is total utility and C is a scaling factor determined at the first iteration, $\Delta U(k)$ = U(k) – U(k-1) and Δt_i = $t_i(k)$ – ti(k-1). Notice that the gradient represents a total differentiation with respect to t_i, including its effect on the demogrant ‘b’. In the uniform tax calculations the program carries out 40 iterations. In the non-uniform calculations 25 separate iterations are carried out for 18 goods. The number of iterations are fixed - no convergence condition has been set when to stop the calculations. In practice, by the 25th iteration the difference between $t_i(25)$ and $t_i(24)$ is almost always below 0.01.[9]

Given that one of the main objectives of the study is to compare results of uniform versus non-uniform taxation, the program calculates total utility and total output under uniform commodity taxation and under conditions when commodity tax rates are individually optimised, and reports the differences in these totals. The difference in total output is:

$$\Delta M = \sum_h W_h \, \Delta \ell_h \qquad (7)$$

The difference in national welfare is converted from utility to monetary values using the Lagrangian multiplier:

$$\Delta U = \sum_h \Delta u_h \,/\lambda. \qquad (8)$$

The difference is reported as a percentage of total output. The Lagrangian is evaluated using the definition:

$$\lambda = \frac{U(R_0 + \Delta R_0) - U(R_0)}{\Delta R_0}$$

[9] Curiously, models involving lower inequality aversion rate tend to show slower convergence in the iterative process.

where R_0 is expenditure on public goods, which is the independent variable in constraint (2a). Changes in R_0 can be used to estimate the marginal utility of public expenditure.

In the baseline model no income tax is specified. In the absence of tax evasion and given zero homogeneous demand in income and prices, a portion of commodity taxes can always be converted into a proportional income tax, without affecting utility or demand. Hence the baseline model is really a commodity cum linear income-tax model and not purely a commodity tax model. To elaborate on this point, suppose that initially a labour input based redistributive model contains only commodity taxes but no income tax. In this situation, the indirect utility function will be:

$$u(\mathbf{p}, w, y) = u(\mathbf{1} + \mathbf{t_0}, W, b) \quad (9)$$

where $\mathbf{t_0}$ are the initial vector of commodity taxes. Now, let us reduce some commodity taxes and convert them into income tax, T. Suppose that this conversion is done by providing a uniform reduction to all gross prices at the rate r, which is offset by a proportional income tax with a constant marginal income tax 'r'. In order to ensure the same proportional changes of all the terms in (9), the demogrant 'b' also has to be reduced by 'r'. After these adjustments $u = u[(\mathbf{1} + \mathbf{t_0})(1 - r), W(1 - r), b(1 - r)]$, which yields exactly the same outcomes as u in (9), because of zero homogeneous utility and demand in prices and earning parameters. Following these changes, the commodity tax rates will be: $t_{i1} = (1 + t_{i0})(1 - r) - 1$ and the linear income tax function will be: $T = rW\ell - rb$. In fact, there are an infinite number of possible linear income tax functions and commodity tax rates defined by different r's that will lead to identical outcomes. In the particular case when r = 0, we have only commodity taxes in the model, which brings us back to our earlier assertion that the baseline model is actually a commodity cum linear income tax model and not purely a commodity tax model.

From equations (1) and (3), an easy way to define the social welfare function is to let all $a_h = 1$, which implies that U is represented by the sum of LES utilities. But of course, a broader range of scenarios can be examined by letting the a_h's to represent different political value judgments. Here we shall follow the inequality aversion approach, based on declining marginal utilities of income. Define the LES utility of taxpayer i as u_i with corresponding marginal utility of income u_{mi}, the social welfare function is U and the inequality aversion rate is z. Now, a preliminary run with a selected uniform commodity tax rate yielded initial estimates for marginal utilities, denoted u_{mi0}. Taking the inverse of u_{mi0} and multiplying it by a scaling factor 'c' to sum up to 15, for the number of taxpayers involved, we obtain social welfare function weights, $g_i = c/u_{mio}$. These weights are approximately inversely related to the marginal utility of income of each taxpayer according to LES. Now, let us define the social welfare function as:

$$U = \sum_i u_i z + u_i(1- z)c/u_{mi0} \quad (10)$$

According to this definition, when z = 1 then we are back to LES utility. But when z = 0, then the social welfare function will have nearly constant weighted marginal utilities of income (cu_{mi}/u_{mi0}), describing a situation where egalitarian objectives are weak or absent. Most of the numerical results discussed in this paper pertain to the original LES utility (z = 1), but in some cases a lower inequality aversion rate (0.3) will be also examined.

It should be noted that in static redistributive models the optimal tax burden depends not only on the inequality aversion rate but also on the average compensated elasticity of labour supply. The optimal average tax rate is negatively related to the average compensated elasticity of labour supply (see Stern (1976), Tuomala (1984), Revesz (1989), Saez (2001)). While this item is part of the model (see eq. (5)), apart from setting various values for β_L, no attempt has been made here to study systematically the effect of this endogenous variable on optimal tax rates.

In order to explain better the numerical results, I developed an approximate formula for optimal commodity tax rates titled the modified inverse elasticity rule. While this formula was not used in the iterative calculations based on (6), it provides a convenient analytical framework to interpret some of the numerical results. It can be extended further to accommodate factors such as administration, compliance, evasion, externalities and leisure complements and substitutes. The mathematical development of the modified inverse elasticity rule is described in appendix 2.

The relevant formula for the baseline model is (A.10 in appendix 2):

$$t_i \approx \frac{1 + t_i - \bar{u}_{mi}\left(1 + t_i + \bar{\varepsilon}_i(t_i - t_{iA})\right)}{\bar{\varepsilon}_i(u_0)} \qquad (11)$$

The term

$$\bar{u}_{mi} = \frac{\partial u / \partial b}{\partial \bar{u}_i / \partial y} \qquad (12)$$

is called the marginal utility ratio of product i. $\partial \bar{u}_i / \partial y$ is the average marginal utility of product i, that is the average marginal utility of income of the consumers of product i, weighted according to their consumption shares (see (A.12) in appendix 2). $\partial u / \partial b$ represents the utility value of the demogrant. Given the same demogrant for all taxpayers, it is the simple average of the marginal utilities of income, that is, the sum of marginal utilities divided by the size of the population. t_{iA} is the average commodity tax rate on products other than i. $\bar{\varepsilon}_i$ is the sales-weighted average price elasticity of good i, and $\bar{\varepsilon}_i(u_0)$ is the corresponding average compensated elasticity of demand.

While the modified inverse elasticity rule is not particularly accurate (as illustrated in appendix 2), it has some advantages over traditional first order conditions. It provides an explicit formula for optimal tax rates. In addition, the unspecified Lagrangian has been replaced by a ratio of marginal utilities $\bar{u}_{mi}$ in (12), which will be of considerable help in some explanations. A similar approximation for optimal tax rates has been worked out by Sandmo (1975), using an entirely different mathematical approach. Sandmo derived his approximation under the assumption of demand separability, i.e.: $\partial q_j / \partial p_k = 0$ for $j \neq k$ No such restrictive assumption was used in developing the modified inverse elasticity rule and the two approximations are not exactly the same.

4. The basic redistributive model

In this section we shall examine the basic model. Various “real life” complexities such as tax evasion, administrative and compliance costs, externalities, paternalistic concerns and non-separability between commodities and leisure will be introduced later. The parameters used in all the scenarios reported in this paper are arbitrary rather than empirical. This approach is perhaps acceptable given that the paper deals with the indirect tax uniformity debate and other broad issues, rather than with the determination of actual tax rates. In any event, empirical estimates can be entered into the source program of the model (located in http://john1revesz.com), which I offer to the interested readers for experimentation and research.

Table 1: Optimal tax rates under the baseline model

	dispersed parameters			bi-polar paramets	inequality aversion=0.3
good	marginal utility ratios	compensated elasticity	tax rate	tax rate	tax rate
1	1.16	-0.39	0.98	0.91	0.32
2	1.29	-0.71	0.85	0.87	0.30
3	0.92	-0.74	0.22	0.92	0.11
4	1.43	-0.84	0.98	0.86	0.32
5	1.13	-1.18	0.48	0.86	0.21
6	1.39	-0.79	0.98	0.87	0.32
7	1.09	-0.90	0.45	0.90	0.21
8	1.23	-0.82	0.68	0.87	0.27
9	1.08	-0.85	0.43	0.87	0.21
10	2.49	-1.28	1.49	1.32	0.39
11	2.29	-0.84	2.69	1.32	0.45
12	2.44	-1.20	1.61	1.32	0.40
13	2.46	-1.24	1.55	1.32	0.39
14	2.58	-1.45	1.32	1.32	0.37
15	2.34	-0.95	2.12	1.32	0.43
16	2.32	-0.91	2.34	1.32	0.44
17	2.58	-1.47	1.31	1.33	0.37
18	2.42	-1.14	1.70	1.32	0.41
Average tax rate on the 9 necessities			0.70	0.89	0.27
Average tax rate on the 9 luxuries			1.83	1.32	0.41
% average tax from expenditure - non-uniform			48.4	51.0	23.8
% average tax from expenditure - uniform			50.8	54.4	25.7
% change in welfare terms compared to uniform solution			3.5	0.2	0.2
% change in total output compared to uniform solution			1.7	-0.3	0.9

Table 1 shows some numerical results from the baseline model where income tax is absent. In the light of the discussion on (9), it represents a particular form of linear income tax combined with commodity taxation. All the calculations were done subject to a fixed public goods expenditure requirement of 10% of total output in the pre-tax situation. In the first two scenarios utilities are as defined by LES without any transformation. Utility transformation, defined by eq. (10), occurs in the third scenario, where the inequality aversion rate is set to 0.3. Looking at the tax rates, the most striking feature is that optimal tax rates are highly differentiated and progressive. This is even true when the inequality aversion rate is reduced from 1 to 0.3. It is not difficult to see from the figures that optimal tax rates are positively related to the marginal utility ratios of goods. Given the definition of these ratios in (12), that means that optimal tax rates are negatively related to the average social marginal utilities of products.

Notice that apart from commodity 3, all other marginal utility ratios are above one. The reason is that the mean of marginal utilities, where each marginal utility of income is counted the same (that is, $\partial u / \partial b$), will usually be larger than the average marginal utility of goods, $\partial\bar{u}/\partial y_i$, because in the later the low marginal utility of incomes associated with higher consumption have a larger weight. That is true even for necessities, where higher income earners tend to consume more, unless the Engel curve is backward sloping (inferior good). With commodity 3 there is a sudden drop in demand at the border between the two consumer groups, which results in low consumption at high income levels (inferior good), leading to the marginal utility ratio of the product falling below one. The finding from the simulations that apart from inferior goods the marginal utility ratio is above one, will assume some importance in later discussion.

It is not difficult to see from the figures that optimal tax rates are negatively related to compensated demand elasticities, particularly among luxuries, in line with the modified inverse elasticity rule. The nearly inverse relationship between tax rates and compensated demand elasticities introduces another source of dispersion in the results besides inequality aversion. Notice that the dispersion of tax rates is much larger under dispersed parameter specifications than under bi-polar specifications, where marginal utility ratios and compensated demand elasticities tend to cluster around similar levels for the two main product groups (see Table C.1 in appendix 3). In the 0.3 inequality aversion rate scenario the demand parameters are dispersed, as in the first scenario, yet the dispersion of optimal tax rates is reduced, because generally tax rates are much smaller.

The change in welfare terms compared with the uniform tax solution is defined in (8) and the change in output in (7). Notice that the gains in welfare and output over the uniform solution are much higher under the dispersed parameters scenario than under the other two scenarios. It appears that these gains are positively correlated with the dispersion of optimal tax rates. Gains in welfare and/or output of non-uniform taxation over the uniform solution in excess of 3% of total output appear also in a number of more complex scenarios that will be discussed later. Another point to note is that the average tax rate tends to be slightly lower under non-uniform than under uniform tax rates.

So far we discussed the structure of optimal commodity tax rates under a linear income tax. At this stage the question arises to what extent the pattern of results would be different under non-linear and non-optimal income taxes. We already know from the Atkinson-Stiglitz (1976) theorem that the optimal non-linear income tax (Mirrlees (1971)) should be combined with uniform commodity tax rates. To examine what happens if income tax is not optimal, the computational model contains a piecewise linear income tax schedule defined by five marginal tax rates over five income intervals. The five income intervals cover about three taxpayers each, but this number can vary in each bracket, depending on variations in taxable income as a result of changing labour supply. Obviously the piecewise linear arrangement can approximate a wide range of non-linear income tax schedules.[10] Table 2 shows results with two arbitrarily defined income tax schedules, one progressive the other regressive. The progressive schedule is made up of five tax brackets with marginal income tax rates increasing from zero at the bottom to 40% at the top bracket, each bracket being 10% higher than the preceding one. In the regressive schedule there is rapid climb from zero marginal tax rate at the lowest bracket to 40% at the second bracket followed by 5% decrease in each of the following three brackets. The shape of the regressive schedule resembles the shape of some of the solutions to the Mirrlees (1971) model reported in Tuomala (1984).

[10] The mathematical framework is changed after income tax is introduced. The net wage rate used in (5) will be $w = W(1 - T_m)$ instead of W, where T_m is the marginal income tax rate. The lump-sum income component 'y' will also change, depending on the curvature of the income tax function. These definitions are discussed in Revesz (1986, 1997). We also touch on them in appendix 5.

Table 2: The effect of income tax

Scenario	Progressive schedule: Dispersed parameters	Progressive schedule: Inequality aversn= 0.3	Regressive schedule: Dispersed parameters	Regressive schedule: Inequality aversn= 0.3
Average tax rate on necessities	0.30	-0.10	0.15	-0.16
Average tax rate on luxuries	0.88	-0.10	0.85	-0.03
% change over uniform in welfare terms	2.7	0.0	3.2	1.4
% change over uniform in total output	0.5	0.3	4.6	2.6
Marginal tax rates				
bracket 1	0	0	0	0
bracket 2	0.1	0.1	0.40	0.40
bracket 3	0.2	0.2	0.35	0.35
bracket 4	0.3	0.3	0.30	0.30
bracket 5	0.4	0.4	0.25	0.25

Notice in Table 2 that in three out of the four scenarios commodity tax rates are differentiated and progressive. The exception is progressive income taxation combined with 0.3 inequality aversion, where the solution turns out to be slightly negative and nearly uniform commodity taxes. When the inequality aversion is one, both progressive and regressive income tax schedules yield strongly differentiated and progressive commodity tax rates and there are significant gains over the uniform commodity tax solution.

Results that we do not publish here using the above scenarios with the 9 commodity linear Engel curves model (discussed in Revesz (1997)), suggest that with linear Engel curves demand optimal commodity tax rates will be nearly uniform even when combined with non-linear income tax schedules. This suggests that Deaton's (1979) theorem is almost valid for non-linear income tax schedules as well. But as explained in Revesz (1997), in a many-goods model linear Engel curves for all goods represents nearly homothetic preferences, which is not supported by empirical evidence.

Needless to say, the results presented in Table 2 are preliminary. Nonetheless, these preliminary results suggest a few points. It appears that generally when Engel curves are not linear then with most exogenously given income tax schedules optimal commodity tax rates will be differentiated and progressive, even in the absence of "real life" complexities. However, there are some non-linear income tax schedules which by themselves almost satisfy distributional objectives and where the associated optimal commodity tax rates tend to be small (positive or negative) and nearly uniform. Extensive simulations with the baseline model (which represents a form of linear income tax), does not suggest that such an outcome will occur with linear income tax schedules.

While this numerical study breaks some new grounds, not all the findings presented here are entirely original. Diamond (1975) analytical study, using a production possibilities frontier model rather than variable labour supply, concluded that in the absence of income tax but in the presence of a demogrant optimal commodity taxes will tend to be differentiated and progressive. As explained earlier, in a basic redistributive model (no administration and evasion) zero income tax yields the same demand and utility outcomes as a range of linear income tax functions combined with proportionally adjusted commodity tax rates. Therefore, the conclusion of Diamond (1975) effectively applies to all linear income tax functions. On a similar vein, Atkinson and Stiglitz (1980) concluded that an optimal commodity tax system, when the income tax is linear progressive, will not generally be uniform under weak separability. Among earlier computational studies, I think only Ebrahimi and Heady (1987) contains all the essential ingredients of a fully-fledged redistributive model, including the presence of lump-sum support payments and variable labour supply. The results presented by Ebrahimi and Heady (1987) indicate that under weakly separable utility optimal commodity tax rates will be differentiated, provided Engel curves are not parallel across households. The heterogeneity of Engel curves causes some departure from strict linearity. Even that slight non-linearity of Engel curves caused a perceptible differentiation in optimal tax rates.[11]

5. Administration and evasion of commodity taxes

In this area we shall discuss three factors: compliance costs by consumers, administrative costs by government and the effect of tax evasion on optimal tax rates.[12] We start with compliance costs. These are defined as administrative and other related costs incurred directly by consumers. They are dead-weight costs (C_i) that reduce real output and are a constant portion (c_i) of the tax revenue from good i. Symbolically:

$$C_i = c_i q_i t_i \qquad (13)$$

Taking the indirect utility definition of the social welfare function in (1), then the dead-weight costs defined in (13) can be incorporated into the social welfare function by subtracting them from personal lump-sum incomes 'y'. The extended social welfare function will be:

$$U = \sum_h a_h u_h(\boldsymbol{P},\ W_h, [y_{oh} - \sum_i c_i q_{ih} t_i]) \qquad (14)$$

Using the modified inverse elasticity rule in appendix 2, we derive in eq. (A.16) the following approximation for the impact of compliance costs on optimal tax rates.

$$\Delta\, t_i(c_i) \approx \frac{[1 + t_i(1 + \bar{\varepsilon}_i)]c_i}{\bar{\varepsilon}_i(u_0)} \qquad (15a)$$

Assuming that $\bar{\varepsilon}_i$ and $\bar{\varepsilon}_i(u_0)$ are the same and equal to -1 (see table 3), we can obtain from (15a) the cruder but simpler approximation:[13]

$$\Delta\, t_i(c_i) \approx -c_i \qquad (15b)$$

Having obtained a theoretical approximation, we can now look in Table 3 at some numerical results. These results were obtained by taking the baseline model dispersed parameters and bi-polar specifications and adding a compliance cost parameter of 4% to two goods, one a necessity the other a luxury. The resulting t_i's are then compared with the original tax rates obtained without compliance costs, shown in Table 1.

[11] Their model is similar to the present one, in the sense that they employ linear Engel curves demand that is not identical across all households.

[12] We are not dealing here with a traditional tax evasion model concerned with evasion, frequency of auditing, penalties and probabilities of detection. For a review of this literature refer to Myles (1995) and Hindriks and Myles (2006). We follow more closely the footsteps of Boadway et. al. (1994).

[13] Note the small differences between $\bar{\varepsilon}_i$ and $\bar{\varepsilon}_i(u_0)$ in Table 3. With LES utility the difference between the two is β_i. Given the condition that $\sum_i \beta_i = 1$, then with 18 goods that means that on the average β_i is below 0.06, and so is the difference between $\bar{\varepsilon}_i$ and $\bar{\varepsilon}_i(u_0)$.

Table 3: Optimal tax rates with compliance costs of 4% of the tax collected

Product No. and type	with compliance	without compliance	impact	Demand elasticities $\bar{\varepsilon}_i$	$\bar{\varepsilon}_i(u_0)$
With dispersed parameters					
Good 5 - necessity	0.44	0.48	-0.04	-1.27	-1.18
Good 18 - luxury	1.66	1.70	-0.04	-1.17	-1.14
Bi-polar specifications					
Good 2 - necessity	0.43	0.87	-0.44	-0.59	-0.40
Good 12 - luxury	1.31	1.32	-0.01	-1.35	-1.32

In the dispersed parameter scenario optimal tax rates decrease in line with what is expected according to (15). But the situation is markedly different in the bi-polar scenario. Due to the clustering of tax rates at a similar level in the original scenario (see table C.1 in appendix 3), even the introduction of a small disturbance can cause a large change. This is shown in the case of good 2, where the optimal tax rate decreases by 44%.[14] The solution tries to minimise the dead-weight compliance costs on low income households by reducing the tax rate, which from (13) is proportional to compliance costs. This perplexing result illustrates the point that the modified inverse elasticity rule is an approximation that does not work well in all cases. Overall, the results indicate that an increase in compliance costs will reduce optimal tax rates by a similar amount to the share of these costs from commodity taxes.

An item closely related to compliance is administration (denoted s). In the present model we define compliance to represent costs paid by the taxpayer, while administration is paid by the government. With s_i public revenue will be affected.

Define the total cost of public tax administration as: $S = \sum_h \sum_i q_{ih} t_i s_i$ (16)

Assuming fixed expenditure on public goods (R_0), the net revenue available for redistribution will be:

$$R = \sum_h \sum_i (q_{ih} t_i - q_{ih} t_i s_i) - R_0 \qquad (17)$$

Total real output will be reduced by the dead-weight cost $S = \sum_h \sum_i q_{ih} t_i s_i$

The incidence of administrative costs on consumers or government has a marked effect on the optimal tax rate. The estimated impact of government administration according to the modified inverse elasticity rule (A.20 in appendix 2)

$$\Delta\, t_i(s_i) \approx \frac{s_i \bar{u}_{mi}[(1 + t_i(1 + \bar{\varepsilon}_i)]}{\bar{\varepsilon}_i(u_0)} \qquad (18a)$$

Again, assuming that $\bar{\varepsilon}_i$ and $\bar{\varepsilon}_i(u_0)$ equal -1 (see table 3), we obtain from (18a) the cruder but simpler approximation:

$$\Delta\, t_i(s_i) \approx -s_i \bar{u}_{mi} \qquad (18b)$$

The notable difference between the approximations in (15) and (18) is the multiplication of the administrative cost term in (18) by $\bar{u}_{mi}$, representing the marginal utility ratio of the product. As indicated in the discussion on Table 1, the marginal utility ratios of products are usually above one and with luxuries even above two. Thus the impacts of government administration costs on optimal tax rates are larger than the impacts of identical compliance costs and there is also an increase in commodity tax progressivity. Table 4 presents results for the same products reported in Table 3.

[14] Experimentation with other necessities in the bipolar scenario showed that imposing a 4% cost per unit tax on any necessity caused a disproportionately large reduction in its optimal tax rate.

Table 4: Optimal tax rates with public administration costs of 4% of the tax collected

Product No. and type	with administration	without administration	impact of admin	impact of compliance	Marginal utility ratio
With dispersed parameters					
Good 5 - necessity	0.43	0.48	-0.05	-0.04	1.13
Good 18 - luxury	1.64	1.70	-0.06	-0.04	2.42
Bi-polar specifications					
Good 2 - necessity	0.41	0.87	-0.46	-0.44	1.11
Good 12 - luxury	1.30	1.32	-0.02	-0.01	2.55

Another item of interest is tax evasion. It is assumed that goods have different propensities for evasion depending on observability and marketing arrangements. Generally, goods and services produced by small business are more evasion prone than those produced or marketed through large organisations. We denote this evasion propensity as e_i, representing a fixed portion of taxes evaded on good i. Another important variable is the dead-weight costs of evasion – its share from the amount of tax evaded is denoted d_i. Obviously d_i depends on how costly it is to carry out tax evasion on the particular good. Taking these definitions, the dead-weight cost of evasion (E_i) is given as:

$$E_i = e_i d_i q_i t_i \qquad (19)$$

This expression is very similar to (13) on compliance and (16) on administration. However, there is a crucial difference between administration/compliance costs and evasion. Assuming given producer prices, administration and compliance has no direct effect on consumer prices. On the other hand, in the presence of evasion, on the average only (1 – e_i) portion of the tax is passed on to consumers. Thus the actual average consumer price in the post-tax situation will be 1 + t_i(1 - e_i) instead of 1 + t_i and the effective tax rate will be t_i(1 – e_i). The changes in prices and lump-sum costs due to evasion can be added to (1). As a result the social welfare function will be:

$$U = \sum_h a_h u_h \left(\boldsymbol{P} - \boldsymbol{te},\ W_h,\ [y_{oh} - \textstyle\sum_i e_i d_i q_{ih}(1-e_i)t_i]\right) \qquad (20)$$

In appendix 2 we derive a fairly complicated extension of the modified inverse elasticity rule incorporating evasion. Complications arise because evasion affects both prices and dead-weight costs, and also because it affects both consumers and the funds available for redistribution (R). Because of its complexity we shall not present here the full formula (A.24 in appendix 2), but only a slightly simpler version where d_i has been set to zero.

$$t_i \approx \frac{1 + t_i(1 - e_i) - \bar{u}_{mi}[1 + t_i(1 - e_i) + \bar{\varepsilon}_i(1 - e_i)(t_i(1 - e_i) - t_{iA})]}{\bar{\varepsilon}_i(u_0)(1 - e_i)} \qquad (21)$$

This approximation is a bit cluttered due to the ubiquity of the (1 - e_i) terms that cannot be cancelled out. In any case, (21) enables us a better understanding of some numerical results. An important point to notice in (21) is the presence of the term (1 – e_i) in the denominator. In the numerator the term $1 + t_i(1 - e_i)$, which represent the post-tax price, has a similar role to $1 + t_i$ in the baseline formula in (11). Thus it would appear that the numerator in (21) is similar to that in the baseline modified inverse elasticity rule. This implies that if d_i is zero, in the presence of tax evasion the optimal tax rate will be about (1/(1-e_i) - 1) higher than without it. Numerical results indeed confirmed this expectation. In the absence of dead-weight costs, the solution tries to recapture revenue lost because of evasion by increasing tax rates on the affected goods. A reduction in tax rates due to evasion can only occur because of associated dead-weight costs.

Now let us take a look at some numerical results. Table 5 is structured the same way as Tables 3 and 4, and the iterative calculations follow the same procedure.

Table 5: Optimal tax rates with commodity tax evasion

Evasion rate = 0.4 Dead weight cost = 25% of tax evaded

Product No. and type	with evasion	without evasion	impact
With dispersed parameters			
Good 4 - necessity	1.05	0.98	+0.07
Good 14 - luxury	2.12	1.32	+0.80
Bi-polar specifications			
Good 1 - necessity	-0.01	0.91	-0.92
Good 11 - luxury	2.16	1.32	+0.84

It is interesting to note in these figures the increasing polarisation of tax rates between necessities and luxuries as a result of evasion. The tax rate on one necessity decreases sharply and on the other necessity it increases only slightly. On the other hand, with both luxuries tax rates increase by over (1/(1-0.4) – 1) = 67% compared with the baseline value, in line with what is predicted by (21) without dead-weight costs. It appears that with necessities the solution tries to reduce evasion related dead-weight costs on low income households by reducing the tax rate, which from (19) is directly linked to dead-weight costs. However, the motive to reduce dead-weight costs is much weaker in respect to luxuries, because of the smaller value of these losses in utility terms. [15]

A frequently voiced argument against differentiated taxes is that large disparities between tax rates may induce more evasion. This is an empirical issue beyond the scope of this paper. Nonetheless, the program includes an option that can be used to explore some aspects of this problem. We assume that evasion also has a tax rate dependent linear component related to the difference between the commodity's tax rate and the average rate ($t_i - t_A$). If the difference is positive then we assume that an extra $\bar{e}_i(t_i - t_A)$ percentage of commodity tax will be evaded, in addition to the effect of the constant e_i specified earlier. If ($t_i - t_A$) is negative then presumably no extra evasion will occur. The combined evasion coefficient will be $\hat{e}_i = e_i + \bar{e}_i(t_i - t_A)$. The program allows for $\bar{e}_i$ to be different from e_i. Setting all $\bar{e}_i$ to 0.3 and all e_i to zero, yielded the results shown in Table 6 for two dispersed parameter scenarios, one with inequality aversion 1 and the other with 0.3. The no evasion scenarios used for comparison have been reported in Table 1. It is evident from these figures that under logarithmic utility, tax rate dependent evasion may have a significant effect in reducing the dispersion of tax rates. However under the 0.3 inequality aversion scenario, the effect on dispersion is minimal compared with the baseline scenario. Moreover, the gains in welfare and output over the uniform solution show little change due to $\bar{e}_i$ when the inequality aversion is low.

[15] When considering this puzzling result, it should be remembered that the model assumes that evasion related dead-weight costs are borne directly by the consumers of q_i. The spillover of some of these dead-weight costs to other consumers has not been taken into account in this simple model.

Table 6: Evasion of 30% above the average tax rate

Dead weight cost = 25% of tax evaded

	Inequality aversion = 1		Inequality aversion = 0.3	
	With evasion	No evasion	With evasion	No evasion
Average tax rate on necessities	0.78	0.70	0.27	0.27
Average tax rate on luxuries	1.43	1.83	0.40	0.41
% change over uniform in welfare terms	1.6	3.5	0.1	0.2
% change over uniform in total output	2.5	1.7	0.9	0.9
% evaded from total tax	6.2		1.1	

Needless to say, in order to optimise commodity tax rates one needs to have empirical information on evasion and administration parameters at the level of products or product groups. To the best of my knowledge such information is not yet available. In regard to evasion propensities and dead-weight losses, probably detailed information will not be published in the foreseeable future, because public authorities are not keen to reveal information that might be useful for tax evaders. In any event, the broad picture about evasion propensities is well known. Goods and services passing through large organisations tend to be much less evasion prone than those passing through small business.

6. Externalities

This section examines the impact of externalities on optimal tax rates Externalities can be both negative and positive, Negative externalities include two well-known items – pollution and congestion. Positive externalities include innovation and network externalities as well as some externalities associated with education, culture and public health.[16] For the sake of modelling simplicity we assume that the costs or benefits of externalities can be quantified in monetary terms and that there is a direct proportional relationship between these costs and benefits and the total consumption of externality generating goods. Denoting the imputed value of the externality as E_i and total consumption of an externality generating good as Q_i , then

$$E_i = \mu_i Q_i \qquad (22)$$

where μ_i is the constant Pigovian cost or benefit rate per unit consumption. In this model we also assume that the imputed costs or benefits of externalities enter into the social welfare function through changes in lump-sum incomes. The definition of the social welfare function in this case will be:

$$U = \sum_h a_h u_h \left(\boldsymbol{P}, W_h, \sum_i \left[y_{oh} + \sum_i \ z_{ih}\, \mu_i Q_i\right]\right) \qquad (23)$$

where z_{ih} represents the portion of E_i received by taxpayer h. By definition, the distribution weights add up to one, i.e. : $\sum_h z_{ih} = 1$. Total output decreases or increases by E_i .

In appendix 2 eq. (A.28) we derive a formula for the impact of a single externality on the optimal tax rate of the externality generating good. After taking out the similar $\bar{\varepsilon}_i$ and $\bar{\varepsilon}_i(u_0)$ terms from the numerator and denominator (as was done also earlier), we obtain the following approximate formula for the change in tax due to externalities.

[16] For a literature review on the taxation of externalities refer to Myles and Hindriks (2006). Another review with particular emphasis on environmental levies and the “double dividend” hypothesis is presented by Fullerton and Metcalf (1997).

$$\Delta t_i(\mu_i) \approx -\frac{\mu_i \sum_h z_{ih} \, {\partial u_h}/{\partial y}}{{\partial \bar{u}_i}/{\partial y}} \qquad (24)$$

The term in the denominator is the average marginal utility of product i, that was defined in (12), and more explicitly in appendix 2 (A.12). The term in the numerator is a weighted average of the social marginal utilities of recipients using the z_{ih} weights. If the externality has the same effect on every member of the population, then $z_{ih} = 1/H$ for everyone. Given the definition of the utility of the demogrant in (12) the numerator in (24) becomes $\mu_i \sum_h {1}/{H} \, {\partial u_h}/{\partial y} = \mu_i \, {\partial u}/{\partial b}$ From definitions (12) this implies that in this case:

$$\Delta t_i(\mu_i) \approx -\mu_i \bar{u}_{mi} \qquad (25)$$

In words, the approximate effect of the externality on the optimal tax rate is the Pigovian rate multiplied by the marginal utility ratio of product i. As explained in the discussion about Table 1, unless the product is an inferior good, the marginal utility ratio is above one. This leads us to conclude that if the distribution of the costs or benefits of an externality are shared equally in the population, then the optimal change in the tax rate will be larger than the Pigovian tax or subsidy rate. The average marginal utility of product i in the denominator of (24) suggests that the higher is the average income level of the consumers of the externality generating good the larger will be the change in the optimal tax rate (up or down) due to the externality.

If the z_{ih} are not equal, then provided the higher z_{ih} are concentrated more among low wage earners and provided the marginal utility of income is decreasing with income, then the numerator in (24) will be higher, hence the tax rate change (up or down) will be larger. The opposite argument applies if the z_{ih} tend to be larger among higher wage earners. In summary, the tax or subsidy induced will be larger the more the externality generating good is a luxury and the more the externality affects low income groups.

A similar conclusion is presented by Sandmo (1975), who derived a similar inverse elasticity formula as (24) for pollution externalities.[17] However, Sandmo (1975) did not claim that the optimal pollution tax rate will generally exceed the Pigovian rate, because he did not test numerically his analytical model.

Having established the modified inverse elasticity rule for externalities, we can take a look at some numerical results. Table 7 presents results with constant proportional externality rates of plus or minus 20%. Externality weights refer to the z_{ih} recipient weights discussed earlier. Decreasing weights refer to the situation where the lowest W taxpayer receives five times more externality than the highest W person. Increasing weights refer to the opposite distribution of z_{ih} weights. As expected from the approximate formula in (24) the departure from the Pigovian rates (plus or minus 20%) is larger the more the externality affects low income persons (decreasing z_{ih}). However, the luxury good does not show larger deviations from the Pigovian rate than the necessity, contrary to what is predicted from (24). A striking feature of Table 7 is that actual changes in tax rates tend to be well above the changes predicted according to the modified inverse elasticity rules presented in (24) and (25). Closer inspection of the results revealed that this is due to changes in the demogrant. In the case of negative externalities the demogrant is increased slightly to compensate for the negative effect of the externality on the utility of taxpayers, and vice

[17] Sandmo (1975) uses similar specifications to the present model, including no income tax and a single externality tax.

versa for positive externalities. The changes in the demogrant induced changes in tax rates that exceed the predictions. This is less pronounced when the inequality aversion is low (=0.3) and the impact on optimal tax rates is closer to the Pigovian rate.

Table 7: Optimal tax rates with externalities

Product No. and type	Pigovian rate μ_i	Type of weights (z_i)	Marginal utility ratio	tax with externality	tax without externality	tax difference
With dispersed parameters						
Good 4 - necessity	-0.2	all equal	1.43	1.68	0.97	+0.71
Good 12 - luxury	+0.2	all equal	2.40	0.95	1.61	-0.66
Good 4 - necessity	-0.2	increasing	1.13	1.55	0.97	+0.58
Good 12 - luxury	+0.2	increasing	1.89	1.04	1.61	-0.57
Good 4 - necessity	-0.2	decreasing	1.78	1.77	0.97	+0.80
Good 12 - luxury	+0.2	decreasing	2.91	0.83	1.61	-0.78
Inequality aversion = 0.3						
Good 4 - necessity	-0.2	all equal	1.09	0.72	0.32	+0.40
Good 12 - luxury	+0.2	all equal	1.14	0.00	0.40	-0.40
Good 4 - necessity	-0.2	increasing	1.06	0.68	0.32	+0.36
Good 12 - luxury	+0.2	increasing	1.16	0.02	0.40	-0.38
Good 4 - necessity	-0.2	decreasing	1.24	0.73	0.32	+0.41
Good 12 - luxury	+0.2	decreasing	1.21	-0.04	0.40	-0.44

Turning to the highly publicised subject of environmental taxes, it can be said that the material discussed in this paper tends to support higher environmental taxes. One reason is the finding that tax increases due to negative externalities can exceed by a factor of two or more the Pigovian tax rates. The other reason is connected with the fact that most fossil fuels are marketed through large organisations and are not particularly evasion prone. Also, the administrative and compliance costs of these taxes are relatively low.

Nonetheless, there is room for caution. The model discussed above, as well as Sandmo (1975), does not explicitly include income taxation.[18] In a recent paper Kaplow (2012) argues that in the presence of income taxation the Pareto optimal solution is to tax or subsidise externalities according to Pigovian rates in a redistributive model. This argument follows from Kaplow's (2006) proposition against differentiated commodity taxation. As will be discussed in section 9, there are quite a few practical and theoretical arguments against the Kaplow (2006) proposition. We shall not repeat them here. Suffice is to say that if despite all the counter-arguments, an appropriate income tax schedule combined with uniform commodity taxes can provide an optimal or near optimal solution to the distributional problem in a realistic model, then Kaplow (2012) is right in saying that externality generating goods should be taxed according to Pigovian rates. But this is very much a conditional statement.

[18] As explained in the discussion about (9), in the present model no income tax yields the same results as linear income tax.

7. Paternalistic concerns

This section examines the impact on optimal tax rates of paternalistic concerns. These include a number of taxes and subsidies provided in line with political judgments about what is in the interest of consumers in the long run. Examples include taxes on "sin goods", subsidies for home buying, educational books and software or expenditure on preventative health care. To some extent, the subsidisation of education and health services is also motivated by paternalistic concerns. Needless to say, these concerns provide another reason for differentiated commodity tax rates.

Given that these items are related to the life-long utility of the consumer, imputation of such costs and benefits raises inter-temporal issues. In the present model we assume that in the static context these imputed costs and benefits are directly proportional to the quantity consumed from the selected good. The imputed value to the consumer per unit consumption is denoted γ_i , therefore total paternalistic costs or benefits to taxpayer h equals to: $G_h = \sum_i q_{ih}\gamma_i$ As before, G_h enters into the utility function of consumers through changes in lump-sum income. The modified social welfare function will be:

$$\mathrm{U} = \sum_h a_h u_h \left(\boldsymbol{P}, W_h, \sum_i [y_{oh} + \sum_i q_{ih}\gamma_i]\right) \tag{26}$$

Comparing this expression with the social welfare function involving externalities defined in (23), we notice that the only difference is that the $\sum_i [y_{oh} + \sum_i z_{ih}\mu_i Q_i]$ term of externalities has been replaced by the $\sum_i [y_{oh} + \sum_i q_{ih}\gamma_i]$ term here. Therefore, there is no need to develop another modified inverse elasticity rule. Suffice is to replace the externality term by the paternalistic term. As with externalities, we focus on a single good – q_i. From (24) it follows that in this case:

$$\Delta t_i(\gamma_i) \approx -\frac{\gamma_i \sum_h \frac{q_{ih}}{Q_i}\partial u_h \big/ \partial y}{\partial \bar{u}_i \big/ \partial y} = -\frac{\gamma_i \partial \bar{u}_i / \partial y}{\partial \bar{u}_i \big/ \partial y} = -\gamma_i \tag{27}$$

The simplification follows from the definition of the average marginal utility of product i in appendix 2 eq. (A. 12). Having obtained a simple approximation for the effect of γ_i on the optimal tax rate, we can take a look at some numerical examples in table 8.

Table 8: Optimal tax rates with paternalistic objectives

Product No. and type	Imputed γ_i rate	with paternalistic objective	without such objective	difference
With dispersed parameters				
Good 2 - necessity	-0.2	1.23	0.85	+0.38
Good 14 - luxury	+0.2	1.08	1.32	-0.24
Bi-polar specifications				
Good 6 - necessity	+0.2	0.02	0.87	-0.85
Good 16 – luxury	-0.2	1.52	1.32	+0.20

With one exception, the changes in optimal tax rates are in line with what is predicted from (27). The exception occurs with good 6 in the bi-polar specifications. We have already noted in the discussion on Table 3 that the modified inverse elasticity rule does not work well with necessities under bi-polar specifications.

8. Leisure complements and substitutes

Complementarity with leisure is one of the central issues in the uniform commodity taxation controversy. All the models favouring uniform taxation assume weakly separable utility between commodities and leisure. It is of some interest to examine the pattern of optimal commodity tax rates when weak separability does not apply, and in particular what is the effect of leisure substitution or complementarity on optimal tax rates. This section will try to provide some answers to these questions.

The starting point for our discussion is the derivative of labour supply with respect to a change in the price of good i. As proved in Revesz (1986, 2005), with weakly separable utility that derivative will be:

$$\frac{\partial \ell}{\partial p_i}\big(u\big(\ell, v(\boldsymbol{q})\big)\big) = \frac{\partial q_i}{\partial m}\ell\theta_w^c - \frac{\partial \ell}{\partial y}q_i \qquad (28)$$

where θ_w^c is the compensated elasticity of labour supply with respect to the wage rate, m is gross income, i.e.: m = Wℓ + y and $\partial\ell/\partial y$ is the derivative of labour supply with respect to lump-sum income.[19] Having established a benchmark equation for weakly separability utility, we can now define deviations from it. We define leisure substitution or complementarity (ϵ_i) as the difference between the actual price derivative of labour and the price derivative corresponding to weakly separable utility:

$$\epsilon_i = \frac{\partial \widetilde{\ell}}{\partial p_i} = \frac{\partial \ell}{\partial p_i} - \frac{\partial \ell}{\partial p_i}(u(\ell, v(\boldsymbol{q})) \qquad (29)$$

For leisure substitutes ϵ_i will be positive because they increase labour supply and vice versa for leisure complements. In order to maintain zero homogeneous labour supply as a function of prices and income, the following condition must be met (see Revesz (1986)):

$$\sum_i \epsilon_i\, p_i = 0 \qquad (30)$$

We can use ϵ_i in a modified inverse elasticity calculation to estimate its approximate effect on optimal tax rates. The detailed derivation is presented in appendix 4. The end result is the following approximation from (B.4):

$$\Delta t_i \approx -\frac{\partial \widetilde{\ell}}{\partial \bar{q}_i}\bar{u}_{mi}\bar{W}_i\bar{T}_i' = -\frac{\bar{\epsilon}_i}{\partial \bar{q}_i/\partial t_i}\bar{u}_{mi}\bar{W}_i\,\bar{T}_i' \qquad (31)$$

where $\dfrac{\partial \widetilde{\ell}}{\partial \bar{q}_i} = \dfrac{\partial \widetilde{\ell}/\partial p_i}{\partial \bar{q}_i/\partial p_i}$ and $T' = \sum_i t_i \dfrac{\partial q_i}{\partial y}$ (32)

$\bar{\epsilon}_i$ is the q_i weighted average value of ϵ_i over H taxpayers, $\bar{W}_i$ is the corresponding average wage rate, $\bar{T}_i'$ is the average portion of commodity taxes from the last dollar expenditure of the consumers of product i. $\bar{u}_{mi}$ is the marginal utility ratio of product i defined in (A.12). The approximation in (31) suggests that the tax rate on i will increase (or decrease) depending by how much tax revenue decreased (or increased) as a result of $\partial\widetilde{\ell}/\partial q_i$ change in labour supply.

It appears that in this model the main purpose of higher taxes on leisure complements than substitutes is in boosting tax revenue for redistribution, rather than in directly improving the utility position of those paying the taxes. The modified inverse elasticity rule is derived in this case (as was done also with other factors) on the assumption that the introduction of ϵ_i has no first order effect on the marginal utility of income of taxpayers. Only its impact

[19] With weakly separable utility demand is independent of the composition of m in terms of Wℓ and y (see Revesz 1986). Therefore, $\frac{\partial q_i}{\partial m} = \frac{\partial q_i}{\partial y}$

on tax revenue has first order effect. This perspective on the role of ϵ_i in a redistributive model is markedly different from the original model of Corlett and Hague (1953), who examined a single consumer economy and demonstrated utility improvement for the representative consumer due to higher tax on a leisure complement. Another point to notice on (31) is that luxuries will have larger movements in tax rates due to leisure dependency than necessities. This is because with luxuries $\bar{u}_{mi}$, $\bar{W}_i$ and $\bar{T}_i'$ are all larger.

Table 9: The effect of leisure complements and substitutes on optimal tax rates

good no.	parameter $\frac{\partial \hat{l}}{\partial \bar{q}_i}$	$\bar{W}_i$ average wage	T_i' marginal tax	$\bar{u}_{mi}$ marginal utility	original model tax	modified model tax	actual tax difference	predicted tax difference
1	0.03	17.7	0.59	1.22	0.98	0.28	-0.70	-0.38
2		18.4	0.59	1.28	0.85	1.15	0.30	
3		13.5	0.59	0.96	0.22	0.54	0.32	
4	-0.02	20.2	0.54	1.41	0.98	1.98	0.99	0.34
5		18.1	0.60	1.18	0.48	0.81	0.33	
6		19.8	0.57	1.37	0.98	1.28	0.30	
7	0.01	16.2	0.60	1.11	0.45	0.62	0.17	-0.10
8		18.0	0.57	1.23	0.68	1.00	0.32	
9	-0.04*	7.5	0.59	1.10	0.43	0.81	0.38	0.19
10	0.05	27.8	0.57	2.22	1.49	0.97	-0.52	-1.97
11		26.0	0.64	2.10	2.69	2.62	-0.08	
12		27.6	0.64	2.21	1.61	1.88	0.26	
13	-0.02	28.1	0.64	2.25	1.55	2.84	1.29	0.78
14		28.8	0.64	2.30	1.32	1.63	0.31	
15		26.6	0.64	2.14	2.12	2.25	0.13	
16		26.3	0.64	2.12	2.34	2.40	0.06	
17		28.8	0.64	2.30	1.31	1.62	0.31	
18	-0.14	28.6	0.64	2.29	1.70	6.00	4.30	5.67
Average tax rate on the 9 necessities					0.70	0.78		
Average tax rate on the 9 luxuries					1.83	2.07		
% change in welfare terms compared to uniform solution					3.5	16.9		
% change in total output compared to uniform solution					1.7	18.0		

* The complementarity parameter on product 9 applies only to the preferences of the 8 lower W taxpayers.

Table 9 compares actual tax rates from the numerical study with the rates predicted using the modified inverse elasticity rule in (31). The numerical study is based on a modified form of LES, which incorporates leisure complements and substitutes. It is explained in detail in appendix 4. The original model we compare with is the dispersed parameters baseline scenario in Table 1. Notice that tax increases for leisure complements tend to be larger than tax decreases for leisure substitutes. This bias combined with larger changes involving luxuries, has led to slightly higher progressivity of optimal tax rates compared with the original model. It is also interesting to note the large improvements in welfare and output compared with the uniform tax solution. However, given that the $\frac{\partial \hat{\ell}}{\partial \bar{q}_i}$ parameters are not based on econometric estimates but were arbitrarily chosen, and the ad hoc nature of modified LES used in these calculations (see appendix 4), perhaps one should observe with caution these striking results.

9. The Laroque-Kaplow proposition

In this section we take a critical look on a recent extension of the uniform commodity taxation argument. A proposition formulated by Laroque (2005) and Kaplow (2006) asserts that given any income tax function combined with differentiated commodity tax rates, and given identical preferences and weakly separable utility between commodities and leisure, it is possible to carry out a reform involving the replacement of non-uniform commodity taxes by adjusting the income tax, so that the utility level of all taxpayers will be maintained or improved. The Laroque-Kaplow (LK) proposition is an extension of the Atkinson and Stiglitz (1976) theorem, stating that given weakly separable utility and a Mirrlees (1971) type optimal non-linear income tax function, there is no need for differentiated commodity taxation. The LK proposition extends the Atkinson-Stiglitz theorem to non-optimal income tax functions as well.

At first sight the LK proposition seems to contradict our numerical results, indicating the optimality of differentiated and progressive commodity tax rates in the presence of a linear income tax and a number of non-linear income tax schedules examined in the simulations. But actually this is not the case. What has been examined in this study is the pattern of optimal commodity tax rates in the presence of an exogenously given (fixed) income tax schedule. A simultaneous change in income and commodity taxation and the issue whether subject to an appropriate change in the income tax schedule optimal commodity tax rates should be uniform or otherwise was not raised in our discussion. Moreover, the LK proposition is based on the assumption of identical preferences. The segmented utility framework employed in the present model assumes that preferences are not the same for all households. Given two different models in terms of perspectives and specifications, it can be said that on purely logical grounds the numerical results presented here neither support nor refute the LK proposition.

At this stage, we could leave the discussion on the LK proposition with this inconclusive statement, however, because the LK proposition has been invoked in the broader debate about uniform versus differentiated commodity taxation, a bit more analysis might be appropriate. Thus far the strongest critique against the LK proposition has been presented by Boadway (2010). Among other things, Boadway points out that the LK proposition may be valid only if the income tax schedule can be adjusted in an appropriate manner to compensate for changes in commodity taxation. If because of political-administrative constraints an appropriate adjustment to the income tax schedule is not carried out, then the LK reform may turn out to be welfare reducing. A bit of thinking on the subject reveals that Boadway’s objection does not just relate to some minor constraint on the shape of the income tax schedule, but relates to a more fundamental problem with the LK proposition.

For the purpose of analysing this issue, we look at a simple corollary to the LK proposition, saying that subject to the conditions specified earlier, it is always possible to carry out a reform that involves setting the commodity tax rates to a uniform level accompanied by an appropriate change in income tax that will result in improvement in social welfare. Notice that this outcome is less restrictive than the one specified by

Laroque (2005) and Kaplow (2006), because we do not assume a Pareto improvement for everyone, but only an improvement in aggregate social welfare. The aggregate welfare proposition follows directly from the Atkinson-Stiglitz theorem. This theorem asserts that uniform commodity taxation combined with a non-linear income tax function will be the best possible mathematical solution to the model. The availability of a global maximum for the social utility function implies that given any initial income tax schedule and set of commodity tax rates, a reform can be carried out based on uniform commodity taxes and an appropriate income tax schedule that will ensure higher social welfare then the original configuration. In the polar case that particular combination will be the Atkinson-Stiglitz solution itself, which represents the global maximum. But of course, there could be many other combinations, based on uniform commodity taxes and changes in the income tax schedule in the direction of the Atkinson-Stiglitz solution that will ensure higher social utility than the starting combination.

So far we discussed the Atkinson-Stiglitz solution in abstract terms. But actually we know that the optimal income tax solution in that model will be the non-linear income tax function from the Mirrlees (1971) model. Numerical and analytical results from that model (see Mirrlees (1971), Tuomala (1984) and Revesz (1989)) indicate that in a finite population the Mirrleesian marginal income tax function will be downward sloping at higher income levels, reaching a marginal income tax rate of zero for the top person. Part of the reason for this outcome is that the Mirrlees (1971) model is based on the assumption of perfect competition in labour markets, where wages exactly match productivities. Needless to explain, such an income tax function is not politically acceptable in a world of imperfect competition and information. That implies that the LK reform may sometimes (but not always) steer toward an income tax solution that is politically unacceptable.[20] Consequently, the warning by Boadway (2010) about the possible infeasibility of carrying out the income tax adjustment required according to the LK proposition, could represent in some cases not a minor adjustment difficulty, but a fundamental obstacle in the way of carrying out the income tax reform required for optimally uniform commodity tax rates.

Apart from "real life" complexities such as tax evasion, administrative and compliance costs, externalities and substitution-complementarity with leisure, all of which invalidate the Atkinson-Stiglitz theorem as well as the LK proposition and are covered in this paper, there is another major complication that is not covered here and is connected with the highly abstract nature of the Atkinson-Stiglitz and LK models, which assume identical household characteristics and preferences. In these models (as well as other optimal taxation models) it is assumed that distributional objectives are solved by optimising income and commodity taxes combined with a uniform demogrant. But a uniform demogrant (or negative tax) is seldom applied in practice. Support agencies around the world attempt to economise on limited funds by providing targeted support based on household characteristics such as: age, household composition, disabilities, health, employment, participation in workfare or trainfare programs and other observable or semi-observable characteristics (see Akerlof (1978)). This approach is in line with Mirrlees (1971) proposition that the best solution is based on taxing or subsidising ability (ie. fixed endowments or potential) rather than income. The effectiveness of differentiated targeted support payments to the needy is of central importance in actual tax-transfer

[20] Cooter (1978) showed that in the presence of proportional commodity taxes, the transversality condition in the Mirrlees (1971) model becomes: $\frac{\partial T}{\partial m} - \sum_i t_i \frac{\partial q_i}{\partial m} = 0$ at the endpoints, where T is income tax. This implies negative marginal income tax rate for the top person and an income tax schedule that appears even less politically acceptable.

systems. The assumption of a uniform demogrant in redistributive models can be justified only on grounds of modelling simplicity.

There have been some attempts in the optimal taxation literature to incorporate differentiated support payments into optimal tax models. Deaton and Stern (1986) examined a model where Engel curves are linear and have their intercepts determined by a linear function of household characteristics. Deaton and Stern demonstrate that uniform commodity taxation is optimal only if the lump-sum grants for each demographic group are an optimal function of household characteristics. The computational study of Ebrahimi and Heady (1987) explores this model with numerical examples, focussing on child benefits. They find that if child benefits are not adequately differentiated according to the number and age of children in the household, then optimal commodity tax rates will not be uniform. Generalising from this result, it can be said that since any support system will fall short of the ideal of providing differentiated lump-sum grants based on actual family subsistence needs and ability to work (due to imperfect screening and political-administrative constraints), non-uniform commodity taxation could be used to compensate partly for the inherent shortcomings in the income support system. This issue is ignored in the LK model, which assumes a uniform demogrant and identical household characteristics apart from W.

Another practical objection relates to the LK assumption that distributional objectives could be addressed mainly through appropriate adjustment to the income tax schedule. However, the well-known evasion and administrative difficulties with income taxation suggest that it should have a limited role in an optimal tax mix. This could be another impediment to the practical application of the LK proposition. [21]

Apart from these practical application issues with a highly stylised and simplistic model, there is also a theoretical problem with the LK model connected with Pareto improvement. To prove Pareto improvement the authors assume that the first step in the LK reform can be carried out in such a manner that both utility and labour supply remain constant. If the distortions in commodity tax rates can be eliminated while labour supply remains constant then a Pareto improvement will occur, because the elimination of price distortions will not cause a reduction in labour supply, which would be the only possible drawback to the LK reform in this simple model. Indeed, with weakly separable utility a choice involving constant utility and labour supply is feasible, but that still leaves open the question whether consumers will always choose that combination, or will some of them elect a position involving higher utility and lower labour supply, which might invalidate the LK proposition. This issue is examined in detail in appendix 5. All told, there is a need for further research on the practical applicability of the LK proposition, but given that this subject is not closely connected to the main theme, I prefer not to pursue it here any further.

10. Summary and qualifications

Taking a broad view on the discussion, the major finding is that given non-linear Engel curves optimal commodity tax rates tend to be progressive and highly dispersed under logarithmic utility specifications, but quite strong progressivity and dispersion often

[21] How to improve the balance between income and commodity taxation has not been explored in this study, due to the lack of suitable parameter estimates. However, the indirect-tax7 program could be used to investigate this subject, because it can accept parameter estimates on evasion, administration and compliance costs for both income and commodity taxation.

persists even when the inequality aversion of society is low. Generally speaking, this conclusion applies to most exogenously given income tax schedules. Moreover, the gains in output and welfare of the non-uniform over the uniform solution are quite substantial. These gains can sometimes exceed 3% of total output even in the baseline model without any complexities added. However, we discovered in the numerical simulations at least one example of a non-optimal non-linear (progressive) income tax schedule, where the corresponding optimal commodity tax rates turned out to be slightly negative and nearly uniform, when inequality aversion is low. Yet, while exceptions are possible, the large majority of numerical results point in favour of differentiated and progressive commodity taxation. This is always the case with linear income tax.

Moreover, the introduction into the model of “real life” complexities, such as tax evasion, administration and compliance costs, externalities and leisure complements and substitutes, tends to increase both tax rate dispersion and the progressivity of optimal tax rates. In addition, they make the gains over the uniform solution much larger. The assumption that these complications are independent factors that are neutral in respect to distributional objectives is shown to be false. The numerical results disprove this assumption, and the modified inverse elasticity formulas of these factors contain marginal utility ratios, which reflect distributional considerations. Tax evasion in particular tends to increase indirect tax progressivity. Among other things, we found that due to distributional considerations externality generating goods should be taxed or subsidised well above the Pigovian rates.

We noted that the dispersion of optimal tax rates is considerably reduced if the inequality aversion of society is low, or if the evasion coefficients of commodities depend also on disparities between commodity tax rates. While these are significant qualifications, they may not satisfy all those insisting on commodity tax uniformity. A number of possible objections can be raised against the non-uniformity proposition and we shall deal here with three that could be the more important ones.

First, the arguments presented in this paper conflict with the tax uniformity theorems discovered by Atkinson and Stiglitz (1976) and Deaton (1979). As mentioned earlier, in my view these theorems have little practical relevance, since they are based on strong simplifying assumptions, ignore “real life” complexities and the non-optimality of actual income tax schedules and selective support payments to the needy. The extension of the Atkinson-Stiglitz theorem by Laroque (2005) and Kaplow (2006) has been analysed in section 9 and its limitations were noted.

Another possible objection relates specifically to the non-linear Engel curves in the present model. The segmentation of LES utility has led to sharp distinction between necessities and luxuries, which could be considered unrealistic and might exaggerate the benefits of progressive indirect taxation. A partial answer to this objection is that at a highly disaggregated level, where qualitative differences are also taken into account, most goods are distinctly either necessities or luxuries, based on consumer group income. Only a minority of goods are located in the middle. In any event, there is room for using in future research utility functions other than segmented LES, or dividing LES into more than two segments.

Another possible objection is that the model is built on the assumption that 18 separate tax rates should be determined for individual products or product groups. The critics could argue that this number of tax rates is excessive, given the additional administrative costs involved and increased evasion opportunities opening up with a large number of tax rates. Naturally the answer to this objection depends on empirical estimates of these additional costs, a subject that is not covered here. Yet numerical examples from the present model suggest that additional administrative and evasion problems are not always sufficient reason to eliminate tax rate differentiation. The numerical examples presented in Table 6 suggest that even when the tax evaded amounts to 30% of the difference between the tax rates of luxuries and the average tax rate, and dead-weight costs amount to 25% of the tax evaded, it is still worthwhile to retain a progressive commodity tax structure, even when the inequality aversion of society is low. Recall from earlier discussion that commodity tax evasion by itself is not a reason to reduce tax rates or bring them closer together. A reduction of tax rates is advisable only if the dead-weight costs of evasion are borne mainly by the poor.[22] Moreover, it is not clear that differentiated commodity taxation will increase administrative and compliance costs, if the general orientation is on taxing more heavily the products of large business rather than small business. Nonetheless, administrative cost considerations limit the number of separate tax rates.

All in all, the model yields strong arguments in favour of differentiated and progressive indirect taxation. But of course, this model offers only a step toward a better understanding and further empirical and computational research is needed.

Appendix 1

User's guide to the computer program titled 'indirect-tax7'

This program yielded the numerical results examined in this paper. The user can specify values such as utility function parameters, wage distribution, inequality aversion of society, piecewise linear income tax schedule, tax evasion on income and selected commodities and associated dead-weight losses, administrative and compliance costs related to income and commodity taxation, externalities, substitution and complementarity with leisure and a few other factors. The original program was developed more than 18 years ago and was reported in Revesz (1997). The current version is much broader than the original.

Program installation and running

The program is located in the website http://john1revesz.com . It is written in a programming language called QBasic, which two decades ago was part of Microsoft Windows. The program is written as a "source code" in a Word file, without being converted into machine language, as are most computer programs sold on the market. A special compiler is needed to convert it into machine code. A QBasic compiler for 64 bit machines is available free of charge from the following website: http://www.qb64.net/ Another QBasic compiler is available for US $60 from: http://www.libertybasic.com/index.html

Having a QBasic compiler, the installation is simple. Open the file named qb64 and then Edit and Paste the entire content of the indirect-tax7 Word file into the QBasic screen. Next, click Run and then Start. Less than a minute later the first prompt will appear on the screen. The prompts present a menu of options. After you finished with the prompts, processing will start automatically and will finish in a couple of minutes. The output file, called tax1.txt, will be located in the same folder as the QBasic compiler. For the best view, it should be opened with Windows Notepad.

[22] Also, let us not forget that with some forms of indirect taxation, such as property taxes, it is advisable to have multiple tax rates for equity reasons.

Incidentally, all the library files associated with the compiler should be kept in the same folder as the compiler file. Because the program is a source code in Word format, it could be corrupted accidentally or during data entry. For that reason, it is advisable to keep a couple of copies for backup, preferably in a different folder.

Data entry

Virtually all the independent variables and parameters in the model can be changed by the user. All the numbers in the DATA lines close to the beginning of the source code can be changed. Above the DATA lines are text lines (marked at the left with an asterisk to indicate that it is a comment and not a program line) explaining what the DATA lines refer to. These include initial wage rates, initial lump-sum incomes, utility parameters, portion of income tax or a particular commodity tax evaded, the percentage of dead-weight costs associated with evasion, administrative and compliance costs of income and commodity taxation, five marginal tax rates defining the piecewise linear income tax schedule, targeted lump-sum support grants to the needy, externalities and leisure complements and substitutes.

Usually there are a number of data lines for each item. Those marked with an asterisk at the beginning of the line, are previously used lines that are ignored by the program. Only the unmarked DATA line is currently active. Considerable care must be exercised when entering new numbers. If the number of entries in the DATA line does not correspond to the number of READs specified underneath, then the program will be corrupted without warning. If there are too few entries, the READ instruction will assume that the missing numbers are zero. If there are too many entries, the surplus numbers will be picked up by the following READ instruction relating to a different variable or parameter. As a result, the output from the program will be meaningless, Please check the parameter summary tables appearing at the beginning of the printout to ensure that correct numbers are used by the program.

Further details on how to operate the program are presented in the user 's guide located in the http://john1revesz.com website.

Appendix 2

The modified inverse elasticity rule

In order to explain better the numerical results, we shall develop an approximate formula for optimal commodity tax rates. While this formula was not used in the numerical calculations, it provides a convenient analytical framework to interpret some of the numerical results. It starts with the baseline model and has been developed further to accommodate additional factors.

In the baseline model (analysed in section 4) suppose that the tax rate t_i is increased by a small amount. The increase has three effects. First, from Roy's Lemma, consumers buying q_i will have their utility reduced by

$$\sum_h \partial u_h/\partial t_i = - \sum_h q_{ih} \, {}^{\partial u_h}/_{\partial y} \qquad \text{(A.1)}$$

where we sum up over H taxpayers.[23]

[23] Note, here utility is derived with respect to lump-sum income (y) rather than income (m = Wℓ+ y), but with weakly separable utility the two derivatives are the same (see Revesz (1986)).

Second, given fixed public expenditure requirements, all the additional tax collected will be redistributed through increase in the uniform demogrant ‘b’. Given equal ‘b’ to all H taxpayers, the social marginal utility of b is

$$\frac{\partial U}{\partial b} = \frac{\sum_h \frac{\partial U_h}{\partial y}}{H} \qquad \text{(A.2)}$$

The amount transferred depends on the revenue generated by increased t_i. Deriving government revenue ($R = \sum t_i q_i$) with respect to t_i yields:

$$\frac{\partial R}{\partial t_i} = q_i + \sum_j t_j \frac{\partial q_j}{\partial t_i} \qquad \text{(A.3)}$$

To simplify matters, we convert the sum on the right side into a single expression. The cross derivatives are usually fairly small terms. The own derivative (j = i) can be used to approximate the sum by a single expression. For that purpose, let us look at two polar cases. When all tax rates apart from t_i are zero, then the total revenue derivative reduces to $t_i \partial q_i / \partial t_i$. When all tax rates are the same (t), and provided labour supply does not change, then $\sum_j t_j \frac{\partial q_j}{\partial t_i} = t \sum_j (1+t) \frac{\partial q_j}{\partial p_i} = t \sum_j p_j \frac{\partial q_j}{\partial p_i} = 0$ where q_j represents total demand by H taxpayers. These cases suggest that the total revenue derivative may be approximated by the following expression:

$$\frac{\partial R}{\partial t_i} = q_i + \sum_j t_j \frac{\partial q_j}{\partial t_i} = q_i + \frac{\partial q_i}{\partial t_i}(t_i - t_{iA}) \qquad \text{(A.4)}$$

where t_{iA} is a product specific average indirect tax rate on goods other than q_i. It can be smaller or larger than t_i. For estimating the net change in total tax revenue, the average t_{iA} on marginal expenditure appears more appropriate than on total expenditure. Combining (A.2), (A.3) and (A.4), the net effect on social utility of giving transfers through the demogrant following a change in t_i is defined by the following expression:

$$\frac{\partial U}{\partial b}\frac{\partial R}{\partial t_i} = \frac{\partial U}{\partial b}\sum_h (q_{ih} + \frac{\partial q_{ih}}{\partial p_i}(t_i - t_{iA}))] \qquad \text{(A.5)}$$

The third element in this redistributive model is the excess burden loss due to increased price distortion. The excess burden (D) is represented by the Harberger triangle under the compensated demand curve:

$$D = -\Delta t_i \Delta q_i /2 \approx -\Delta t_i \frac{\partial q_i(u_0)}{\partial t_i}\frac{\Delta t_i}{2} \approx -t_i \frac{\partial q_i(u_0)}{\partial t_i}\frac{t_i}{2} \approx -t_i^2 \frac{\partial q_i(u_0)}{\partial p_i} /2 \qquad \text{(A.6)}$$

We assumed in this approximation that the compensated demand curve is linear, hence its price derivative is constant. It is also assumed that the appropriate Harberger triangle starts from zero tax, thus in (A.6), $\Delta t_i = t_i$. The excess burden loss is borne directly by the taxpayer. Deriving (A.6) with respect to t_i and then multiplying by the consumer’s marginal utility of income term and summing over h taxpayers we obtain:

$$\frac{\partial D}{\partial t_i} \approx -\sum_h \frac{\partial u_h}{\partial y} t_i \frac{\partial q_{ih(u_0)}}{\partial p_i} \qquad \text{(A.7)}$$

At the optimum, the three utility effects, represented by (A.1), (A.5) and (A.7) must add up to zero. Summing up while converting derivatives into elasticities (ε) we obtain:

$$0 \approx -\sum_h q_{ih} \,{}^{\partial u_h}/_{\partial y} + \frac{\partial U}{\partial b}[\sum_h (q_{ih} + \frac{q_{ih}\varepsilon_{ih}}{p_i}(t_i - t_{iA}))] - \sum_h \frac{\partial u_h}{\partial y} t_i \frac{q_{ih}\varepsilon_{ih(u_0)}}{p_i} \qquad \text{(A.8)}$$

Multiplying all the terms in (A.8) by ${}^{p_i}/_{\sum_h q_{ih}}$ we arrive at:

$$0 \approx -\frac{\partial \overline{u_i}}{\partial y} p_i + \frac{\partial u}{\partial b}[p_i + \bar{\varepsilon}_i (t_i - t_{iA})] - t_i \frac{\partial \overline{u_i}}{\partial y}\bar{\varepsilon}_i(u_0)] \qquad \text{(A.9)}$$

where $\frac{\partial \overline{u_i}}{\partial y}$, $\bar{\varepsilon}_i$, and $\bar{\varepsilon}_i(u_0)$ are all weighted averages, with the weights given by the q_{ih} purchases of individual consumers. After dividing all terms by $-\frac{\partial \overline{u_i}}{\partial y}$ and replacing p_i by 1

+ t_i we arrive at the final formula:

$$t_i \approx \frac{1 + t_i - \bar{u}_{mi}\left(1 + t_i + \bar{\varepsilon}_i(t_i - t_{iA})\right)}{\bar{\varepsilon}_i(u_0)} \qquad (A.10)$$

The term
$$\bar{u}_{mi} = \frac{\partial u/_{\partial b}}{\partial \overline{u_i}/_{\partial y}} \qquad (A.11)$$

is called the marginal utility ratio of product i. $\partial u/\partial b$ has been defined in (A.2). $\partial \bar{u}_i/\partial y$ is the average marginal utility of product i. It is defined as:

$$\frac{\partial \overline{u_i}}{\partial y} = \frac{\sum_h \frac{\partial u_h}{\partial y} q_{ih}}{\sum_h q_{ih}} \qquad (A.12)$$

The approximate formula in (A.10) is referred to as the modified inverse elasticity rule.

Table A.1 Comparing tax rates from iterations with predictions from the modified inverse elasticity rule[24]

	Dispersed		Bipolar		Inequality=0.3	
Good	iterations	predictions	iterations	predictions	iterations	predictions
1	0.98	0.83	0.91	0.56	0.32	0.30
2	0.85	0.82	0.87	0.60	0.30	0.29
3	0.22	-0.48	0.91	0.54	0.11	-0.09
4	0.98	1.00	0.86	0.59	0.32	0.32
5	0.48	0.14	0.86	0.61	0.21	0.12
6	0.98	0.97	0.87	0.58	0.32	0.32
7	0.45	0.20	0.90	0.55	0.21	0.13
8	0.68	0.58	0.87	0.57	0.27	0.23
9	0.43	0.21	0.87	0.57	0.21	0.14
10	1.49	2.12	1.32	2.04	0.39	0.39
11	2.69	2.58	1.32	2.07	0.45	0.47
12	1.61	2.18	1.32	1.98	0.40	0.40
13	1.55	2.14	1.32	2.00	0.39	0.40
14	1.32	1.98	1.32	2.09	0.37	0.37
15	2.12	2.50	1.32	2.00	0.43	0.46
16	2.34	2.51	1.32	1.94	0.44	0.46
17	1.31	1.96	1.33	1.98	0.37	0.36
18	1.70	2.26	1.32	1.99	0.41	0.42
Average	0.94	1.08	1.04	1.13	0.31	0.29

Given that (A.10) is used in some analytical explanations, it is of some interest to compare numerical results obtained from this approximation with the results from iterations based on (6). Table A.1 displays such comparisons. While some of the predictions from the modified inverse elasticity rule are substantially different from those of the iterations, it should be noted that in these baseline scenarios the difference in average tax rates between predictions and iterations is less than 14%. Also, the results from

[24] It should be noted that the modified inverse elasticity rule does not predict well when a non-linear income tax is included in the model. It provides reasonable predictions only without income tax. However, as explained in section 3, in the present model no income tax yields the same results as linear income tax.

predictions and iterations are strongly correlated. Despite the differences, the results appear to be sufficiently close to justify the application of the modified inverse elasticity rule for analytical purposes. Moreover, Sandmo (1975) has worked out a similar approximation for optimal tax rates using an entirely different mathematical approach.

Compliance costs

In this paper the main benefit derived from the modified inverse elasticity rule is in providing some estimates on the impact of various factors on optimal tax rates. We start the discussion with compliance costs that are borne by the taxpayer. The definition of these costs in (13) is: $C_i = c_i q_i t_i$ Taking the indirect utility definition of the social welfare function in (1), then the dead-weight costs defined in (13) can be incorporated into the social welfare function by subtracting them from lump-sum income 'y'. The extended social welfare function will be:

$$U = \sum_h a_h u_h(\boldsymbol{P},\ W_h, [y_{oh} - \sum_i c_i q_{ih} t_i]) \qquad \text{(A.13)}$$

Applying the modified inverse elasticity rule to (A.13), we may notice that the presence of c_i has a visible effect only on the personal utility term originally shown in (A.1). Deriving (A.13) with respect to t_i we obtain:

$$\sum_h \partial u_h/\partial t_i = - \sum_h q_{ih} {}^{\partial u_h}/_{\partial y} - \sum_h {}^{\partial u_h}/_{\partial y} \left(c_i q_{ih} + t_i c_i {}^{\partial q_{ih}}/_{\partial t_i} \right) \qquad \text{(A.14)}$$

Combining (A.14) with the unchanged (A.5) and (A.7) terms and performing the operations from (A.5) to (A.9), we arrive at the following extended version of the modified inverse elasticity rule:

$$t_i \approx \frac{1 + t_i + [(1 + t_i) + \bar{\varepsilon}_i t_i] c_i - \bar{u}_{mi}\left(1 + t_i + \bar{\varepsilon}_i (t_i - t_{iA})\right)}{\bar{\varepsilon}_i (u_0)} \qquad (A.15)$$

Subtracting (A.10) from (A.15) yields the estimated impact of c_i on the optimal tax rate :

$$\Delta\, t_i(c_i) \approx \frac{[1 + t_i(1 + \bar{\varepsilon}_i)] c_i}{\bar{\varepsilon}_i (u_0)} \qquad (A.16)$$

Notice that in this subtraction we implicitly assumed that marginal utilities and demand derivatives with and without c_i are the same. This is not strictly correct, but judging from the numerical results, this assumption does not distort by much the approximations.

Administration costs

Similar exercise can be carried out in regard to the impact of s_i on the optimal tax rate. With s_i public revenue will be affected. Assuming fixed expenditure on public goods (R_0), the net revenue available for redistribution will be:

$$R = \sum_h \sum_i (q_{ih} t_i - q_{ih} t_i s_i) - R_0 \qquad \text{(A.17)}$$

Deriving (A.17) with respect to t_i, we obtain a revised transfer term instead of (A.5)

$$\frac{\partial U}{\partial b} \frac{\partial R}{\partial t_i} = \frac{\partial U}{\partial b} [\sum_h (q_{ih} - s_i (q_{ih} + t_i \frac{\partial q_{ih}}{\partial p_i}) + \frac{\partial q_{ih}}{\partial p_i} (t_i - t_{iA}))] \qquad \text{(A.18)}$$

Combining (A.18) with (A.1) and (A.7) and performing the operations from (A.6) to (A.9), we arrive at the following extended version of the modified inverse elasticity rule:

$$t_i \approx \frac{1 + t_i - \bar{u}_{mi}\left(1 + t_i - s_i\, [(1 + t_i) + \bar{\varepsilon}_i t_i] + \bar{\varepsilon}_i (t_i - t_{iA})\right)}{\bar{\varepsilon}_i (u_0)} \qquad (A.19)$$

Subtracting (A.10) from (A.19) yields the estimated impact of s_i on the optimal tax rate :

$$\Delta\, t_i(s_i) \approx \frac{s_i \bar{u}_{mi} [(1 + t_i) + \bar{\varepsilon}_i t_i]}{\bar{\varepsilon}_i (u_0)} \qquad (A.20)$$

Tax evasion

We can apply a similar procedure to assess the impact of tax evasion (e_i) on optimal tax rates. Unfortunately, in this case the analysis becomes quite complicated, because in addition to dead-weigth costs ($E_i = e_i d_i q_i t_i$) the average price also changes due to evasion. The average price will be: $p_i = 1 + t_i(1 - e_i)$ with the price derivative being $dp_i/dt_i = 1 - e_i$ instead of one, as in earlier calculations. The effect of the change in the average price has a pervasive effect on all relevant calculations. Unlike with administration and compliance costs (as well as externalities that will be discussed later) we cannot derive a simple estimate for the impact of tax evasion on optimal tax rates. Nonetheless we shall work out here the modified inverse elasticity rule with tax evasion, because it will help us to explain some perplexing results from the iteration.

From (20) the social welfare function with evasion is given as:

$$U = \sum_h a_h u_h \left(\mathbf{1} + \boldsymbol{t}(\mathbf{1} - \boldsymbol{e}),\ W_h,\ [y_{oh} - \sum_i t_i e_i d_i q_{ih}(1 - e_i)]\right)$$

Deriving consumer utility by t_i (corresponding to (A.1)) we obtain:

$$\sum_h \frac{\partial u_h}{\partial t_i} = \sum_h \frac{\partial u_h}{\partial p_i}\frac{\partial p_i}{\partial t_i} = \sum_h \frac{\partial u_h}{\partial p_i}(1 - e_i) =$$

$$= -\sum_h [q_{ih}\frac{\partial u_h}{\partial y}(1 - e_i) + e_i q_{ih} d_i (1 - e_i) - d_i e_i \frac{\partial q_{ih}}{\partial p_i}(1 - e_i)^2] \qquad \text{(A.21)}$$

Deriving public revenue (corresponding to (A.5) we obtain

$$\frac{\partial R}{\partial t_i} = \frac{\partial U}{\partial b}\sum_h [q_{ih}(1 - e_i) + \frac{\partial q_{ih}}{\partial p_i}(1 - e_i)(t_i(1 - e_i) - t_{iA})] \qquad \text{(A.22)}$$

Finally we come to the third element, the excess burden. The Harberger triangle will be:

$$D = \frac{\partial u}{\partial y}\Delta q_i \frac{\Delta t_i}{2} = \frac{\partial u}{\partial y} t_i (1 - e_i) t_i \frac{\partial q_i}{\partial p_i}(u_0)\frac{(1 - e_i)}{2} = \frac{\partial u}{\partial y}\frac{((1 - e_i)t_i)^2}{2}\frac{\partial q_i}{\partial p_i}(u_0)(1 - e_i)$$

Therefore:

$$\frac{\partial D}{\partial t_i} = \frac{\partial u}{\partial y}(1 - e_i) t_i \varepsilon_i(u_0)\frac{q_i}{p_i}(1 - e_i) = \frac{\partial u}{\partial y}(1 - e_i)^2 t_i \varepsilon_i(u_0)\frac{q_i}{1 + t_i\ (1 - e_i)} \qquad \text{(A.23)}$$

Assembling (A.21), (A.22) and (A.23) and following the steps from (A.5) to (A.9), we arrive at the modified inverse elasticity rule with tax evasion.

$$t_i \approx \frac{1 + t_i(1 - e_i) - e_i d_i (1 + (1 - e_i) t_i) - e_i d_i (1 - e_i)\, t_i \bar{\varepsilon_i}}{\bar{\varepsilon_i}(u_0)(1 - e_i)} -$$

$$- \frac{\bar{u}_{mi}[1 + t_i(1 - e_i) + \bar{\varepsilon_i}(1 - e_i)(t_i(1 - e_i) - t_{iA})]}{\bar{\varepsilon_i}(u_0)(1 - e_i)} \qquad (A.24)$$

This is a complicated expression that is difficult to use in order to assess the impact of e_i on the optimal tax rate. In the text we use (A.24) with the dead-weight loss (d_i) set to zero, in order to explore certain features of the solution.

Externalities

The starting point for deriving the formula for externalities is the extended social welfare function defined by (23) in section 6. In order not to complicate the analysis with cross-price effects, we shall concentrate here only on a single externality. In this case he extended social welfare function will be:

$$U = \sum_h a_h u_h(\boldsymbol{P},\ W_h, [y_{oh} + \sum_h z_h \mu_i \sum_h q_{ih}]) \qquad \text{(A.25)}$$

Applying the modified inverse elasticity rule to (A.25), by deriving it with respect to t_i we obtain:

$$\sum_h \partial u_h/\partial t_i = -\sum_h q_{ih}\ {}^{\partial u_h}/_{\partial y} + \mu_i \sum_h z_h\ {}^{\partial u_h}/_{\partial y}\ (\sum_h {}^{\partial q_{ih}}/_{\partial t_i}) \qquad \text{(A.26)}$$

Notice that the externality term multiplying μ_i is opposite in sign to the utility derivative based on Roy's Lemma. Combining (A.26) with the unchanged (A.5) and (A.7) terms and performing the operations from (A.5) to (A.9), we arrive at the following extended version of the modified inverse elasticity rule:

$$t_i \approx \frac{1 + t_i - \bar{\varepsilon}_i\mu_i \frac{\sum_h z_h \, {\partial u_h}/{\partial y}}{{\partial \bar{u}_i}/{\partial y}} - \bar{u}_{mi}\left(1 + t_i + \bar{\varepsilon}_i(t_i - t_{iA})\right)}{\bar{\varepsilon}_i(u_0)} \qquad (A.27)$$

Subtracting (A.10) from (A.27) yields the estimated impact of μ_i on the optimal tax rate :

$$\Delta\, t_i(\mu_i) \approx \; - \frac{\bar{\varepsilon}_i\mu_i \sum_h z_h \, {\partial u_h}/{\partial y}}{\bar{\varepsilon}_i(u_0) \, {\partial \bar{u}_i}/{\partial y}} \qquad (A.28)$$

The numerator contains the Pigovian term multiplied by the weighted marginal utilities of income of externality recipients. The denominator is the average marginal utility of product i defined in (A.12).

Appendix 3

Table C.1: Further data pertaining to Table 1 in the text

	bipolar – split factor = 0.99			dispersed - inequality aversion = 0.3		
good	marginal utility ratio	compensated elasticity	tax rate	marginal utility ratio	compensated elasticity	tax rate
1	1.05	-0.18	0.91	1.07	-0.44	0.32
2	1.11	-0.40	0.87	1.12	-0.74	0.30
3	1.05	-0.19	0.92	0.98	-0.72	0.11
4	1.13	-0.48	0.86	1.15	-0.86	0.32
5	1.13	-0.45	0.86	1.09	-1.19	0.21
6	1.12	-0.43	0.87	1.14	-0.82	0.32
7	1.10	-0.38	0.90	1.07	-0.90	0.21
8	1.11	-0.41	0.87	1.11	-0.82	0.27
9	1.12	-0.43	0.87	1.06	-0.84	0.21
10	2.55	-1.28	1.32	1.27	-1.14	0.39
11	2.55	-1.27	1.32	1.26	-0.91	0.45
12	2.55	-1.32	1.32	1.27	-1.10	0.40
13	2.55	-1.31	1.32	1.27	-1.12	0.39
14	2.55	-1.25	1.32	1.28	-1.24	0.37
15	2.55	-1.31	1.32	1.27	-0.95	0.43
16	2.55	-1.35	1.32	1.27	-0.94	0.44
17	2.55	-1.32	1.33	1.28	-1.25	0.37
18	2.55	-1.32	1.32	1.27	-1.06	0.41

Appendix 4

The mathematical framework with leisure complements and substitutes

This section examines two separate issues. First, we derive the modified inverse elasticity approximation for leisure complements and substitutes presented in eq. (31) in the text. Following this, we outline the modified LES utility function that was used to obtain the numerical results reported in Table 9.

In regard to the modified inverse elasticity rule, we may notice that among the three components analysed in appendix 2, only the second component, that is tax revenue, is visibly affected by the presence of $\epsilon_i' s$. Roy's Lemma and the excess burden term are not affected in any obvious way. For the purpose of obtaining a simple expression for the revenue term, we find it useful to redefine ϵ_i from (29) as:

$$\epsilon_i = \frac{\partial \tilde{\ell}}{\partial p_i} = \frac{\partial \tilde{\ell}}{\partial q_i}\frac{\partial q_i}{\partial p_i} \qquad \text{(B.1)}$$

Now, the revised revenue term (based on (A.5)) will be:

$$\frac{\partial U}{\partial b}\frac{\partial R}{\partial t_i} = \frac{\partial U}{\partial b}\sum_h (q_{ih} + W_h T_h' \frac{\partial \tilde{\ell}}{\partial q_i}\frac{\partial q_i}{\partial p_i} + \frac{\partial q_{ih}}{\partial p_i}(t_i - t_{iA}))] \qquad \text{(B.2)}$$

W_h is the wage rate of taxpayer h and T_h' is the marginal indirect tax rate of consumer h. Note, this model excludes income tax, so here the marginal tax rate refers to the share of commodity taxes in the last dollar of expenditure. From these definitions it is evident that the term added to (B.2) simply refers to the change in labour income multiplied by the marginal indirect tax rate, yielding the change in tax revenue due to the ϵ_i factor.

Combining (B.2) with the unchanged (A.1) and (A.7) terms and performing the operations from (A.5) to (A.9), we arrive at the following extended version of the modified inverse elasticity rule:

$$t_i \approx \frac{1+t_i - \bar{u}_{mi}\left(1+t_i+\frac{\partial \tilde{\ell}}{\partial \bar{q}_i}\bar{\varepsilon}_i \bar{W}_i \overline{T_i'} + \bar{\varepsilon}_i(t_i - t_{iA})\right)}{\bar{\varepsilon}_i(u_0)} \qquad \text{(B.3)}$$

Subtracting (A.10) from (B.3) and cancelling the similar price elasticity terms in the numerator and denominator, we obtain:

$$\Delta t_i \approx -\frac{\partial \tilde{\ell}}{\partial \bar{q}_i}\bar{u}_{mi}\bar{T}_i'\bar{W}_i \qquad \text{(B.4)}$$

where all terms on the right hand side represent q_i weighted average values.

Modified LES

To obtain the numerical results in Table 9, we used a modified version of LES With modified LES labour supply in (5) is given as:

$$\ell = Z - q_L = Z(1-\beta_L) - \alpha_L - \frac{\beta_L}{w}\left(y - \sum_j p_j \alpha_j\right) - \sum_i k_i q_i \qquad \text{(B.5)}$$

where the k_is are constants and $-\sum_i k_i q_i$ is added as an extra term.
This corresponds to the redefinition of utility in (3) as:

$$u = \sum_i \beta_i \log(q_i - \alpha_i) + \beta_L \log(q_{L0} + \sum_i k_i q_i - \alpha_L) \qquad \text{(B.6)}$$

where q_{L0} is the value of leisure under weakly separable utility, that is in the absence of the $\sum_i k_i q_i$ term. Notice that from definitions (B.1) and (B.5), $\frac{\partial \tilde{\ell}}{\partial q_i} = - k_i$

Obviously, the k_is introduce an inter-dependency between commodities and leisure. It is not clear whether (B.6) can be solved to yield explicit global formulas for commodity demand and labour supply. But even without global formulas, it is possible to obtain local formulae in the neighbourhood of a given vector $\mathbf{q_{io}}$. In our numerical study $\mathbf{q_{io}}$ and q_{L0} are the optimal quantities obtained from the weakly separable utility model. Suppose the k_is satisfy the following initial condition:

$$\sum_i k_i\, q_{i0} = 0 \tag{B.7}$$

From (B.7) it follows that (B.5) will continue to satisfy zero homogeneous labour supply at $\mathbf{q_{io}}$. Moreover, the consumer will be able to make the same choices as under weakly separable utility. But actually the consumer will not choose the previous basket of goods and leisure. Given condition (B.5) his/her new choice of leisure will be:

$$q_L = q_{L0} + \sum_i k_i\,(q_i - q_{i0}) \tag{B.8}$$

Given different labour supply (ℓ = Z - q_L), the consumer's income will change. To balance the budget constraint the net change in expenditure will be:

$$W(\ell - \ell_0) = W\Delta\ell = \sum_i p_i\,(q_i - q_{i0}) = \sum_i p_i\,\Delta q_i \tag{B.9}$$

To determine Δq_i, we employed the marginal consumption propensities $\partial q_i/\partial y$.
From the LES demand equation in (4): $\partial q_i/\partial y = \beta_i\,/p_i$ (B.10)

After taking out β_L and inflating the $\beta_i s$ of commodities by 1/ (1 - β_L) to equal one, we obtain from (B.10):

$$\sum_i \frac{\partial q_i}{\partial y} p_i = \sum_i \frac{\widehat{\beta_i}}{p_i} p_i = \sum_i \widehat{\beta_i} = 1 \tag{B.11}$$

where $\widehat{\beta_i}$ are the recalibrated values of β_i for commodities. To maintain proportionality between $\partial q_i/\partial y$ and Δq_i and ensure that the budget constraint (B.9) is maintained:

$$\Delta q_i = W(\ell - \ell_0)\frac{\widehat{\beta_i}}{p_i} = W(\ell - \ell_0)\frac{\widehat{\beta_i}}{1+ t_i} \tag{B.12}$$

Adding the values found for $\Delta\ell$ from (B.8) and Δq_i from (B.12) to the original q_{L0} and $\boldsymbol{q_{i0}}$ variables, yields new values for commodities and leisure that can be used to evaluate the utility in (B.6). From then on, the calculations follow the procedure described in section 3.

Although the redefinition of LES presented here is fairly ad hoc, it should be noted that the resulting demand system perfectly satisfies the budget constraint and almost perfectly satisfies zero homogeneity of demand and labour supply near the point $\mathbf{q_{io}}$. This approach was adopted in order to enable numerical testing of the impact of leisure complements and substitutes, without having to write a new program. Arguably, non-separable indirect utility functions, such as those used in the computational studies of Ebrahimi and Heady (1987) and Murty and Ray (1987) are better suited for testing the effect of leisure non-separability on optimal tax rates, but that challenge is left for future research.

Appendix 5

Constant utility and labour supply in the LK proposition

A crucial assumption in the LK proposition, which is used to demonstrate Pareto improvement, is that with weakly separable utility it is possible to reduce commodity tax rates to zero and by an appropriate increase in income tax, to reach a situation where labour supply and utility remain the same as before the reform. Putting aside the possible imperfect adjustability of income tax noted by Boadway (2010), and assuming that appropriate changes can be carried out, the question arises whether following such tax changes will all taxpayers actually choose the same utility and labour supply as before? The proofs on this point presented by Laroque (2005) and Kaplow (2006) are fairly opaque, and I shall discuss here my interpretation on them.

From the definition of weakly separable utility as U = f(v(c),ℓ), where v is a sub-utility function of commodities only, it is not difficult to see that following variations in earning parameters and prices that leave v constant, provided ℓ remains constant so will U. Hence, constant ℓ and U is a feasible outcome. However there are another two possible outcomes associated with constant v changes – either increase ℓ and reduce U or decrease ℓ and increase U. It is not difficult to see that the option to increase ℓ and reduce U will not be a rational choice. That still leaves open either constant ℓ and U or decreasing ℓ and increasing U.

The possible choice of decreasing ℓ and increasing U was not considered by Laroque (2005) and Kaplow (2006), who assumed that following price-income changes that leave v constant, it would be rational for all taxpayers to maintain both ℓ and U constant, which would subsequently lead to higher U when the reduction in price distortions is taken into account. But in fact there are constant v variations where it would be quite rational, for at least some taxpayers, to reduce ℓ and thereby increase U. These taxpayers would eventually benefit from more leisure and less distorted prices. It all depends on changes in the real values of the post income-tax earning parameters – the net wage rate (w) and "virtual" lump-sum income (y). For discussion about these post income-tax earning parameters and the "virtual budget" framework refer to Revesz (1986, 1989), Roberts (2000) and Saez (2001).

Briefly, the net wage rate is defined as: $w = W(1 - T')$ (D.1)

and "virtual" lump-sum income as: $y = T'W\ell - T$ (D.2)

T represents total income tax and T' is the marginal income tax rate. Labour supply function is given as $\ell = \ell(w, y, \mathbf{p})$. Because w and y are the principal determinants of labour supply, given changes in these net earning parameters the assumption about constant utility and labour supply associated with changes involving constant v may not necessarily be true.

My earlier studies (Revesz (1986, 1997)) suggest that progressive commodity taxation will lead to higher labour supply than equal revenue generating progressive income taxation. This is an important issue, because according to computational results, given non-linear Engel curves, inequality aversion and no income tax, the optimal commodity tax structure will be progressive. If following the replacement of progressive commodity taxation by (presumably progressive) income taxation, labour supply of at least a section of the population is reduced, then government revenue might be curtailed following the LK reform, which could lead to a fall in the real value of the demogrant. The effect on aggregate social welfare will depend on the social marginal utilities of winners and losers.

The fact that with weakly separable utility, each consumer has the option to keep v, U and ℓ constant, seems to be behind Laroque's (2005) argument that since in the post-reform situation "*the agents have access to exactly the same menu (v'(Y), Y) as before, they choose the same labour supply*". Y is defined as before tax income, Y= Wℓ. Where Laroque (2005) probably made a mistake is in assuming that having access to the same menu (expressed in terms of Wℓ) will ensure the same labour choice, despite possible changes in the net earning parameters w and y, which are ignored in his analysis.

Kaplow (2006) mathematical proof is more elaborate than that of Laroque (2005), which makes it easier to identify a possible mistake. Kaplow (2006) claims in Lemma 1 that given weakly separable utility, it is possible to construct an intermediate income tax function (denoted $T^0(W\ell)$), which will ensure that following the replacement of non-uniform commodity taxes by $T^0(W\ell)$, utility and labour supply (ℓ) remain unchanged. Kaplow (2006, p. 1240) defines $T^0(W\ell)$ so that:

$$v(\mathbf{t}, T, W\ell) = v(\mathbf{t^*}, T^0, W\ell) \quad \text{for all } W\ell. \qquad \text{(D.3)}$$

While this definition is valid for constant v, what Kaplow (2006) did not recognise is that having a different income tax function (T^0) will affect labour supply. From definitions (D.1) and (D.2) following the replacement of T by T^0 labour supply is given as $\ell(w, y, \mathbf{p}) = \ell(W(1 - T^{0\prime}), (T^{0\prime}W\ell - T^0), \mathbf{p})$ This implies that labour supply as a function of W will remain the same as before the reform only if the following implicit differential equation is satisfied :

$$\ell_0(W) = \ell(T^{0\prime}, T^0, \mathbf{p}) \qquad \text{(D.4)}$$

where $T^{0\prime}$ is the derivative of T^0 with respect to $W\ell$ and ℓ_0 is pre-reform labour supply. The likely mistake in Lemma 1 (p. 1241) occurs in the second equation where Kaplow claims that:

$$U(v(\mathbf{t}, T, W\ell), \ell) = U(v(\mathbf{t^*}, T^0, W\ell), \ell) \text{ for all } W\ell. \qquad \text{(D.5)}$$

Kaplow explains that this equality follows from (D.3). In (D.5) the left hand side refers to the pre-reform situation and the right hand side to the post-reform situation. As explained earlier, in (D.5) ℓ on the left hand side would equal ℓ on the right hand side only if the differential equation in (D.4) is satisfied. But there is no reason to believe that generally the same $T^0(W\ell)$ function will satisfy both conditions (D.3) and (D.4), which leads us to conclude that ℓ on the two sides of (D.5) is not the same, in contradiction to what Kaplow set out to prove. In my view, the mathematical proofs presented by Laroque (2005) and Kaplow (2006) require further scrutiny.

Appendix 6

Some other numerical results

In this section we review two other sets of numerical results. The discussion is cursory. The objective is to familiarise the reader with other options available with this computational model and their possible role in future explorations of the optimal tax mix. Other options are described briefly in the user's guide located in the website: http://john1revesz.com

All the scenarios presented in this paper are based on the assumption that public goods expenditure requirement (R_0) amounts to 10% of total output in the no-tax zero demogrant situation.[25] Table E.1 presents some results showing the effect of changing R_0. Evidently, the higher is R_0 the higher will be the average tax burden. The dispersion of tax rates (in proportional terms) is slightly reduced with increasing R_0.

[25] Given that in the no-tax situation total output is 20-25% higher than in the post-tax situation, R_0 as a percentage of actual output in most simulations is over 12% rather than 10%.

Table E.1: The effect of changing public goods expenditure requirements

	Dispersed parameters		Inequality aversion=0.3	
% public expenditure from total no-tax output	20	0	20	0
Average tax rate on necessities	0.83	0.60	0.32	0.22
Average tax rate on luxuries	2.00	1.67	0.46	0.38
% change over uniform in welfare terms	3.9	3.2	0.2	0.2
% change over uniform in total output	1.5	4.7	0.7	1.1

Another item incorporated into the program is the provision of targeted grants to the needy. This subject has been already discussed in section 9. In the scenarios described in Table E.2 targeted support is included by assigning a lump-sum payment amounting to around one third of average income to the three lowest W taxpayers ($200) and a payment amounting to half that size ($100) to the fourth taxpayer from the bottom. These payments remain constant regardless of the value of the demogrant 'b' found in the course of optimisation. These lump-sum grants are supposed to represent selective support payments that are determined on the basis of considerations outside the model.

Table E.2: The effect of targeted support grants to the needy

	Dispersed parameters		Inequality aversion = 0.3	
Scenario	With targeted grants	Without grants	With targeted grants	Without grants
Average tax rate on necessities	0.71	0.70	0.22	0.27
Average tax rate on luxuries	1.78	1.83	0.38	0.41
% change over uniform in welfare terms	3.2	3.5	0.2	0.2
% change over uniform in total output	1.9	1.7	0.8	0.9
Total output in $	11408	11358	13652	13521

The figures in Table E.2 show that introduction of targeted support grants does not alter much the tax rates in the dispersed parameter scenario, but causes a larger reduction of tax rates in the 0.3 inequality aversion rate scenario. The bigger change occurs in total output. With the introduction of a total of $700 targeted grants at the expense of 'b', total output in the first scenario increases by $50 and in the second by $131. These are quite substantial gains. There has been little computational modelling done so far on the pros and cons of targeted support to the needy compared to a uniform demogrant, although such modelling could be useful in the policy debate about negative income tax versus targeted assistance.

References

Akerlof, G. A. (1978), The economics of 'tagging' as applied to the optimal income tax, welfare programs and manpower planning, *American Economic Review*, **68**, *8-19.*

Alm, J. (1996), What is an 'optimal' tax system?, *National Tax Journal*, **49/1**, 117-133.

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 (1976), *Lectures in Public Economics*, New York, McGraw-Hill.

Bastani, S., Blomquist, S., Pirttila, J. (2013), *How should commodities be taxed? A*

counterargument to the recommendation in the Mirrlees Review, Working Paper No. 4240, Centre for Economic Studies & Ifo Institute, Ludwig Maximilians University
Boadway, R. (2010), Efficiency and redistribution: an evaluative review of Louis Kaplow's The Theory of Taxation and Public Economics, *Journal of Economic Literature*, **48:4**, 964-979.
Boadway, R. Marchand, M. and P. Pestieau (1994), Towards a theory of the direct-indirect tax mix, *Journal of Public Economics*, **55**, 71-88.
Boadway, R. and P. Pestieau (2003), Indirect taxation and redistribution: the scope of the Atkinson-Stiglitz theorem, in *Imperfect Economics: Essays in Honor of Joseph Stiglitz*, edited by R. Kanbur and R. Arnott, The MIT Press, 387-403.
Christiansen, V. A. (1984), Which commodity tax rates should supplement the income tax? *Journal of Public Economics*, **24**, 195-220.
Cooter, R. (1978), Optimal tax schedules and rates: Mirrlees and Ramsey, *The American Economic Review*, **68,** *756-768.*
*Corlett, W. C. and D. C. Hague (1*953), Complementarity and the excess burden of taxation, *Review of Economic Studies*, **21,** 21-30.
Cremer, H. and F. Gahvari (1995), Uncertainty, optimal taxation and the direct versus indirect tax controversy", *The Economic Journal*, **105**, 1165-1179.
Cremer, H., Pestieau, P. and J. Rochet (2001), Direct versus indirect taxation: the design of tax structure revisited, *International Economic Review*, **42**, 781-799.
Deaton, A. (1979), Optimally Uniform Commodity Taxes, *Economic Letters*, **2**, 357-361.
Deaton.A. and N. Stern (1986), Optimally uniform commodity taxes, taste differences and lump-sum grants, *Economic Letters*, 20, pp. 265-266.
Diamond, P. A. (1975), A many person Ramsey tax rule, *Journal of Public Economics*, **4**, 335-342.
Ebrahimi, A. and Heady C. (1987) Tax design and household composition, *Economic Journal*, **98**, 83-96.
Feldstein, M. S. (1972), Distributional equity and the optimal structure of public prices, *The American Economic Review*, **62**, 32-36.
Fullerton, D. and Metcalf, G. E. (1997), *Environmental taxes and the double-dividend hypothesis: Did you really expect something for nothing?*, NBER Working Paper 6199, www.nber.org/papers/w6199.
Hindriks, J. and Myles G. D. (2006), *Intermediate Public Economics*, London, The MIT Press.
Kaplow, L. (2006), On the undesirability of commodity taxation even when income taxation is not optimal, *Journal of Public Economics*, **90**, 1235-1250.
Kaplow, L. (2012), Optimal control of externalities in the presence of income taxation, *International Economic Review*, **53,** pp. 487-509.
Laroque, G. (2005), Indirect taxation is superfluous under separability and taste homogeneity: a simple proof, *Economic Letters*, **87**, 141-144.
Mirrlees, J. A. (1971), An exploration in the theory of optimal taxation, *Review of Economic Studies*, **38**, 179-208.
Mirrlees, J., S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (2011), *Tax by Design: the Mirrlees Review*, Oxford University Press.
Murty, M. N. and Ray, R. (1987), Sensitivity of optimal commodity taxes to relaxing leisure/goods separability and to the wage rate, *Economic Letters*, **24**, 273-277.
Myles, G. D. (1995), *Public Economics*, Cambridge University Press.

Naito, H. (1999), Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency, *Journal of Public Economics*, **71**, 365-388.

Nygard, O. E (2008) *Optimal Indirect Taxation: A Review of Theoretical and Empirical Studies,* Statistics Norway – Research Department, http://www.ssb.no/emner/05/90/doc_200819/doc_200819.pdf

Ray, R. (1986), Sensitivity of 'optimal' commodity tax rates to alternative demand functional forms, *Journal of Public Economics*, **31**, 253-268.

Revesz, J. T. (1986), On some advantages of progressive indirect taxation, *Public Finance/Finances Publiques*, **41**, 182-197

Revesz, J. T. (1989), The optimal taxation of labour income, *Public Finance/Finances Publiques*, **44**, 433-475.

Revesz, J. T. (1997), Uniform versus non-uniform indirect taxation: some numerical results, *Public Finance/ Finances Publiques,* **52**, 210-234.

Revesz, J. T. (2005), The optimal tax mix reconsidered, *Public Finance/ Finances Publiques*, **1-2/54**, 1-21.

Roberts, K. (2000), A Reconsideration of the optimal income tax in Hammond, P. J. and Myles, G. D. (eds.), *Incentives and Organisation: Papers in Honour of Sir James Mirrlees*, Oxford University Press.

Saez, E. (2001), Using elasticities to derive optimal income tax rates, *Review of Economic Studies*, **68**, 205-229.

Saez, E. (2002), The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes, *Journal of Public Economics*, **83**, 217-230.

Sandmo, A. (1975), Optimal taxation in the presence of externalities, *Swedish Journal of Economics*, **77**, pp. 86-98.

Srinivasan, P. V. (1989), Redistributive impact of 'optimal' commodity taxes, *Economic Letters*, **30**, 385-388.

Stern, N. H. (1976), On the specifications of models of optimum income taxation, *Journal of Public Economics*, **9**, 123-162.

Stiglitz, J. E. (1982), Self-selection and Pareto efficient taxation, *Journal of Public Economics*, **17**, 213-240.

Thomas, R. L. (1987), *Applied Demand Analysis*, UK., Longman Group.

Tuomala, M. (1984), On the optimal income tax, *Journal of Public Economics*, 23, pp. 351-366.