Some contributions to the theory of optimal indirect taxation

The purpose of this book is to contribute to the static theory of optimal indirect taxation, by examining a number of analytical and computational models. The discussion focuses mainly on the controversy whether indirect taxes should be uniform or otherwise. It concentrates mainly on optimisation models, be it analytic or numerical, though attention is also given to alternative approaches for looking at tax reforms. The principal argument in this book is that contrary to the widely accepted view in support of uniform indirect taxation, based on two highly abstract mathematical models, under real world conditions optimal indirect tax rates should be differentiated and progressive.

I have a BA degree in Economics from Tel-Aviv and a Master of Commerce from the University of N.S.W. From the mid 1980s until my retirement a few years ago, I worked as a research economist in the Australian Public Service. Over the last 32 years, I published seven articles on optimal taxation in refereed journals.

John Revesz

Some contributions to the theory of optimal indirect taxation

A literature review and some computational results
John Revesz

Some contributions to the theory of optimal indirect taxation
Some contributions to the theory of optimal indirect taxation

A literature review and some computational results

John Revesz
Imprint
Any brand names and product names mentioned in this book are subject to trademark, brand or patent protection and are trademarks or registered trademarks of their respective holders. The use of brand names, product names, common names, trade names, product descriptions etc. even without a particular marking in this work is in no way to be construed to mean that such names may be regarded as unrestricted in respect of trademark and brand protection legislation and could thus be used by anyone.

Cover image: www.ingimage.com

Publisher:
Editorial Académica Española
is a trademark of
International Book Market Service Ltd., member of OmniScriptum Publishing Group
17 Meldrum Street, Beau Bassin 71504, Mauritius
Printed at: see last page

Copyright © John Revesz
Copyright © 2018 International Book Market Service Ltd., member of OmniScriptum Publishing Group
Contents

1. Introduction and synopsis........................................................................................................3

2. A literature review..................................................................................................................5
   2.1 Introduction to chapter 2.................................................................................................5
   2.2 Theoretical considerations...............................................................................................9
      2.2.1 The Ramsey model and its extensions......................................................................9
      2.2.2 The uniform taxation theorems..............................................................................10
         2.2.2.1 Non-linear income tax and the Atkinson-Stiglitz theorem..............................10
         2.2.2.2 The main objections raised against the Atkinson-Stiglitz theorem.............11
         2.2.2.3 Non-distributional factors.............................................................................13
         2.2.2.4 Other possible objections to tax uniformity.................................................14
         2.2.2.5 The self-selection approach..........................................................................14
         2.2.2.6 The Laroque-Kaplow proposition: sub-optimal income tax.......................15
         2.2.2.7 Optimal linear income tax and the Deaton theorem.....................................16
   2.3 Empirical-computational studies......................................................................................17
      2.3.1 Methodological issues.............................................................................................17
      2.3.2 Representative consumer models.........................................................................19
      2.3.3 Many-consumer economy.....................................................................................20
         2.3.3.1 Models without a lump-sum grant.................................................................20
         2.3.3.2 Models with sub-optimal lump-sum grants.................................................22
         2.3.3.3 Models with endogenously determined optimal lump-sum grants...........23
      2.3.4 Relevance to the uniform taxation theorems.........................................................24
      2.3.5 Some critical remarks on flexible demand functions.............................................25
      2.3.6 An alternative approach: Optimal marginal reforms.........................................25
   2.4 The treatment of income tax............................................................................................26
   2.5 Uniformity and policy related studies .............................................................................27
   2.6 Conclusions.....................................................................................................................28

Appendices to chapter 2
   2.1 Joint optimisation of Mirrleesian income tax with indirect taxation.........................29
   2.2 Joint optimisation of piecewise-linear income tax with indirect taxation................31


1 Keywords: Optimal taxation, commodity taxation, indirect taxation, uniform taxation, externalities, tax policy.
3. A computational model........................................................................................................35
   3.1 Introduction to chapter 3.................................................................................................35
   3.2 Some recent policy related studies..................................................................................36
   3.3 The mathematical framework of the segmented LES model..........................................38
   3.4 The redistributive model.................................................................................................42
      3.4.1 The model without income tax..................................................................................42
      3.4.2 The model with a non-linear income tax.....................................................................44
      3.4.3 The two composite goods model...............................................................................46
      3.4.4 Analytical arguments in favour of progressive indirect taxation...............................48
      3.4.5 Discussion on the Laroque-Kaplow proposition.......................................................50
      3.4.6 Subsidies for education and health.............................................................................54
   3.5 Other factors justifying differentiated indirect taxation..................................................54
      3.5.1 Compliance costs.......................................................................................................55
      3.5.2 Government administration costs...............................................................................56
      3.5.3 Tax evasion.................................................................................................................57
      3.5.4 Externalities...............................................................................................................60
      3.5.5 Paternalistic concerns................................................................................................63
      3.5.6 Leisure complements and substitutes.........................................................................64
   3.6 Summary and qualifications.............................................................................................67
      3.6.1 The main results........................................................................................................67
      3.6.2 Some qualifications.....................................................................................................67
      3.6.3 Policy issues...............................................................................................................68

   Appendices to chapter 3
   3.1 User’s guide to the computer program titled ‘indirect-tax8’.............................................69
   3.2 The modified inverse elasticity rule..................................................................................75
   3.3 Further details about the setup of the numerical model.....................................................80
   3.4 Supplementary tables.......................................................................................................84
   3.5 The mathematical framework with leisure complements and substitutes.........................85

References....................................................................................................................................88
Post-script ................................................................................................................................94
A correction to the discussion on the Atkinson-Stiglitz theorem.............................................94
1. Introduction and synopsis

The purpose of this book is to contribute to the static theory of optimal indirect taxation, by examining a number of analytical and numerical models. The discussion focuses mainly on the controversy whether indirect taxes should be uniform or otherwise. It concentrates mainly on optimisation models, be it analytic or numerical, though attention will be also given to alternative approaches for looking at tax reforms. The book usually refers to indirect taxation rather than commodity taxation, which is the more commonly used term in the optimal taxation literature. The aim is to indicate that the discussion covers also taxes on services and property taxes.

Chapter 2 is divided into two main parts. The first part reviews the analytical-mathematical literature, the second part is concerned with empirical-computational studies. The analytical review covers all major theoretical studies belonging to the Ramsey and Mirrlees traditions. A critical review of these mathematical models reveals that from an empirical point of view, analytical arguments in favor of uniform indirect taxation seem rather weak and unrealistic, hence determining the optimal tax structure remains an empirical issue.

However, reviewing the empirical-computational studies published so far, reveals that most of them operate under rather restrictive and simplistic frameworks. These studies provide little computational support for uniformity, particularly when the models approach real-world complexity. It appears that in a many-consumer economy, differentiated and progressive indirect taxation is likely to be the optimal solution.

Attention is also given to some recently published policy-related studies, which advocate uniform indirect taxation, despite the fact that this position is supported only by a few highly abstract models, and is contradicted when real-world complexities are taken into account.

The two appendices to chapter 2 explain how to carry out numerical joint optimisation of direct and indirect taxes with a Mirrleesian income tax function, and how to carry out joint optimisation with a piecewise-linear income tax.

Chapter 3 substantiates the main arguments presented in chapter 2, by presenting numerical results from a computational model, based on segmented linear expenditure system (LES) demand. This model throws some light on the structure of optimal commodity taxes in a many-person many-goods static setting. One of the major findings is that with non-linear Engel curves and linear income tax, optimal commodity tax rates will be progressive and highly dispersed under logarithmic utility specifications. The dispersion of optimal tax rates is reduced, if the inequality aversion rate of society is low. When an exogenously given non-optimal and non-linear income tax schedule is included in the model, usually there is still a need for differentiated and progressive indirect taxation. These findings are in marked contrast to the continuing preoccupation of much of the literature with uniform indirect taxation for distributional purposes.

The results also indicate that if tax evasion incurs substantial dead-weight costs (say, above 25%), it usually reduces optimal tax rates by over a half of the evasion/revenue ratio of the product, with the reduction being larger for necessities and smaller for luxuries.
Private compliance costs and government administration costs reduce optimal tax rates by a similar percentage to the share of these costs from taxes.

In a model with linear income tax, the effect of externalities on optimal tax rates substantially exceeds the corresponding Pigovian tax rates or subsidies.

The main benefit of higher taxes on leisure complements than leisure substitutes, appears to be in boosting tax revenue for redistribution, rather than in improving 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 factors tend to reinforce the progressivity of optimal commodity tax rates.

For the purpose of analysing the numerical results, I develop in the appendix to chapter 3 an approximate formula for optimal tax rates, called the modified inverse elasticity rule. This rule and its extensions, provide a simple analytical tool to predict approximately optimal indirect tax rates. The modified inverse elasticity rule is not applied in this book to calculate optimal tax rates, which are found using iterative calculations.

The appendix also contains a user’s guide for the computer program that was employed to obtain the numerical results presented in the book.

Finally, I discuss some economically advisable directions for tax reform in EU countries. I suggest that an advisable direction for reform would be to replace part of income tax by higher taxes on non-necessities produced and/or marketed by large organisations, and also to raise property taxes on luxury housing. There seems to be considerable scope for welfare improving and evasion reducing reforms. No consideration is given here to the political feasibility of such reforms. Overall, the book provides strong arguments in favour of differentiated and progressive indirect taxation.

Chapter 2 is an expanded version of an article titled “A literature review on optimal indirect taxation and the uniformity debate” that was published in the journal Hacienda Publica Espanola/ Review of Public Economics (2016/3), 218: pp. 107-138. Chapter 3 is an expanded version of an article titled “A numerical model of optimal differentiated indirect taxation” that was also published in the journal Hacienda Publica Espanola/ Review of Public Economics (2014/4), 211: pp. 9-66. Both chapters contain much more material than the original publications.

\[\text{2 The co-author of this article preferred not to participate in this book.}\]
2. A literature review *

2.1 Introduction to chapter 2

A large part of the static optimal taxation literature is concerned with optimal indirect taxation, i.e. taxes on the supply and demand for different goods, focusing on the indirect tax structure and the optimal mix between direct and indirect taxes. I refer in this book to indirect taxation rather than commodity taxation, which is the more commonly used term in the optimal taxation literature. The aim is to indicate that the discussion covers also taxes on services as well as property taxes. The discussion excludes border tariffs, because this subject more appropriately belongs to the field of international trade. The discussion is restricted mainly to optimisation models, be it analytical or numerical.\(^3\) The merits of another numerical approach for examining the tax uniformity issue will be examined in section 2.3.6.

Since in a static model any uniform indirect tax structure can be replaced by a proportional direct (income) tax, the tax-mix issue and the issue of optimal indirect taxes are closely related.\(^4\) We shall examine in this chapter the optimal tax-mix in some detail, with some new results presented in appendices 2.1 and 2.2. In context of analyzing the optimal tax-mix, the question arises: is there a need for differentiated indirect taxation? The purpose of this chapter is to provide an up-to-date review of the literature on this topic.

In recent years several policy studies have appeared, arguing that indirect taxes should generally be uniform and distributional concerns should be left solely to direct taxes and welfare benefits (Mirrlees et al (2011); Arnold et al. (2011); European Commission (2013); IMF (2014); NOU (2014)). At the same time, there is a tendency to recommend a shift from direct to indirect taxes.\(^5\) In light of this, this review may be relevant to the ongoing tax policy discussion.

A recent literature review on optimal indirect taxation is presented by Tuomala (2016), as part of a broader survey on redistributive taxation. Tuomala places a strong emphasis in his survey on mathematical analysis and various extensions to the Mirrlees model. An earlier review is presented by Ray (1997) on optimal commodity taxes and optimal reforms. The present review differs substantially from both Tuomala (2016) and Ray (1997) in terms of scope and methodology.\(^6\) Like Santoro (2002), who surveys the marginal reform approach (see section 2.3.6), this survey is narrower in scope, since it focuses mainly on the tax uniformity issue. For that purpose, I put more attention to empirical-computational studies, their results, methods,

\(^3\) The reason is that the uniform taxation theorems arise from mathematical optimisation models, therefore it seems logical to frame objections against them through optimisation models.

\(^4\) Such interchangeability between taxes is only valid in the absence of heterogeneous evasion and administrative costs. When discussing the optimal tax mix, I could not avoid touching on certain issues in the theory of optimal income taxation.

\(^5\) The principal arguments are that taxes on consumption affect less saving and investment decisions compared to income taxes, induce less work disincentives, are less vulnerable to tax evasion, and hence are more favorable to economic growth. In this paper I assume that labour income and capital income are taxed separately, so we do not have to enter into intertemporal problems associated with capital taxation.

\(^6\) Ray (1997) places a strong emphasis in the survey on his own studies (see sections 2.3.3.1 and 2.3.3.2).
and modelling assumptions. The review of analytical results is used mainly to explain computational studies and help in the interpretation of their results. Thus, I do not go into formal mathematical derivations, but focus instead on the intuition and assumptions behind the tax structure results, with special attention to the uniformity issue.\(^7\)

The first attempt to analyse the optimal tax problem was Ramsey (1927). He posed the question: Suppose the government needs to raise a certain amount of tax revenue, how should it collect this revenue in a way that will minimize welfare losses? This started what could be called the 'Ramsey tradition', covering models where proportional commodity taxes are combined with zero or linear income tax. Ramsey studied a single-consumer economy in which direct taxation is absent. Later, the model was extended to the many-consumer case (Diamond, 1975; Diamond and Mirrlees, 1971), with the possibility of linear income tax (Deaton, 1979a; Deaton and Stern, 1986; Boadway and Song, 2016), and with externalities (Sandmo, 1975). The model was also interpreted and studied by several other scholars (Baumol and Bradford, 1970; Besley and Jewitt, 1995; Corlett and Hague, 1953; Deaton, 1979b, 1981, 1983; Dixit, 1970, 1975, 1979; Feldstein, 1972; Lerner, 1970; Mirrlees, 1974; Munk, 1977; 1980; Sadka, 1977; Samuelson, 1982; Sandmo, 1976; Stiglitz and Dasgupta, 1971). The only significant contributions in the Ramsey tradition that suggested uniform commodity tax rates in a many-person setting are Deaton (1979a) and Deaton and Stern (1986), for linear Engel curves and linear income tax.

James Mirrlees gave birth to what may be called the 'Mirrlees tradition' (Mirrlees, 1971, 1976), covering many-person models with an optimal non-linear income tax. Here the basic problem is asymmetric information, since the tax authority cannot discern each individual's ability, but can observe only income. Lump-sum taxation of abilities would be the first-best solution, but it is infeasible in practice. Mirrlees (1971) studied originally non-linear income tax, but later studies have focused on a mix, where the optimal or improved non-linear income tax is combined with non-linear or proportional indirect taxes (Atkinson and Stiglitz, 1976; Mirrlees, 1976; Cooter, 1978; Stiglitz, 1982; Laroque, 2005; Kaplow, 2006; Kanbur et al., 2006)). As with the Ramsey tradition, several theoretical extensions exist, such as allowing for heterogeneous preferences (Mirrlees, 1976; Saez, 2002a), heterogeneous endowments (Mirrlees, 1976; Cremer et al., 2001; Boadway and Pestieau, 2003), limited mobility between skills (Naito, 1999), incorporating externalities (Pirttilä and Tuomala, 1997), minimising the deprivation index (Kanbur et al., 2006), and non-separable utility between commodities and leisure (Christiansen, 1984; Edwards et al., 1994; Nava et al., 1996; Jacobs and Badoway, 2014). A number of studies in the ‘Mirrlees tradition’ suggested uniform commodity taxation in a many-person setting. These include: Atkinson and Stiglitz (1976), Stiglitz (1982), Laroque (2005) and Kaplow (2006).\(^8\)

The basic many-person optimal tax problem for \(H\) individuals, covered in the uniform taxation theorems of Atkinson and Stiglitz (1976) and Deaton (1979a), can be put as follows:

---

\(^7\) Several surveys deal with the traditional theory of optimal commodity taxation, for example Auerbach (1985).

\(^8\) Yet, other studies in the Mirrlees tradition objected to tax uniformity, for reasons other than non-separable utility. These include: Pirttilä and Tuomala (1997), Naito (1999), Cremer et al. (2001), Boadway and Pestieau (2003), Kanbur et al. (2006) and a couple of sections in Mirrlees (1976).
\[
\text{Max } \sum_{i=1}^{N_h} W(u(t_1, \ldots, t_N), T(m_{h})) \tag{1.1}
\]
\[
s.t. \sum_{h=1}^{N} \sum_{i=1}^{N} t_i p_i x_{ih} + \sum_{h=1}^{N} T(m_h) = G + Hb \tag{1.2}
\]

where \( W(.) \) is the welfare function of individual utilities \( u(.) \). \((t_1, \ldots, t_N)\) are the \( N \) commodity tax rates; \( T(m_h) \) is the income tax function; \( \omega_h \) is the unobservable ability level (or long-term wage rate) of taxpayer \( h \); \( p_i \) is the constant producer price of commodity \( i \); \( x_{hi} \) represents the quantities demanded by individuals; \( G \) denotes fixed government expenditure on public goods; and \( b \) is a uniform lump sum grant to \( H \) consumers. Gross income is: \( m_h = \omega_h \ell_h \) where \( \ell \) is labour supply. Labour is the only primary input in these static models. The objective is to maximize the welfare function in (1.1) subject to the tax revenue constraint in (1.2). This is a joint optimisation problem of income and commodity taxes. Generally, this problem can be solved by maximizing with respect to indirect tax rates, the income tax function, and the lump-sum grant. In the Atkinson-Stiglitz (briefly A-S) model, the search is for the fully optimal differentiable non-linear income tax function. In the Deaton model only linear income taxes are considered. In empirical-computational models, the cardinalist welfare transformation, \( W(.) \), which determines the social valuation weights, is usually defined by a single parameter, called the inequality aversion rate. Econometric techniques are used to provide estimates for utility-demand parameters in these models.

The mathematical correctness of the A-S and Deaton uniform taxation theorems is not in doubt. The debate is on the empirical validity of these highly abstract models. Briefly, the uniform taxation theorems are based on the following assumptions:

- Identical preferences for all consumers along with different wage rates (abilities).
- Weakly separable utility between commodities and leisure
- No administrative and compliance costs and no tax evasion
- The absence of other complicating factors such as externalities, merit goods and imperfect competition.
- Income redistribution is carried out through a uniform lump-sum grant for everyone.
- The A-S model takes it for granted that the optimal non-linear income tax schedule can be implemented in practice.
- The Deaton theorem assumes linear Engel curves for all goods.

Numerous objections have been raised against these assumptions, including:

- Preferences are not identical due to different compositions of households, different needs and endowments apart from different abilities.
- Weakly separable utility between commodities and leisure is probably not valid in respect to many goods.
- Complicating factors, such as administrative costs, tax evasion, externalities and merit goods, all violate the uniform taxation theorems.
- The optimal non-linear income tax schedule a la Mirrlees (1971), which is part of the A-S model, is often not politically acceptable.
- Linear Engel curves for all goods, which is postulated in Deaton’s theorem, is not supported by empirical evidence.
- Income redistribution is much more complicated than the provision of a uniform lump-sum grant for everyone. In actual support systems there may be non-optimal elements.
A number of empirical-computational studies assume zero or sub-optimal lump-sum grants to represent inadequate redistribution. This modelling approach seems to be more relevant to the situation in developing countries.

The counter-arguments to uniformity appear in both the theoretical literature and in empirical-computational studies. Most theoretical contributions relevant to the uniformity debate deal with models in the Mirrlees tradition, while all empirical-computational studies, except for one, deal with the extended Ramsey framework, and usually ignore income tax. The advantages of empirical-computational models compared with purely deductive mathematical studies, are in that they can much more easily handle complexities, and can provide some quantitative appreciation of the magnitudes involved. On the other hand, they can only describe optimal outcomes in particular situations, and are not well suited to derive general principles that are applicable under a broad range of circumstances. Moreover, some empirical-computational models deal with a single person economy and are largely irrelevant for distributional considerations. All relevant computational models assume constant producer prices. In principle, supply side could be also brought in, and general equilibrium models could be applied, yet no such optimisation studies exist to my knowledge.9

The present review indicates that the arguments in favor of a uniform tax structure seem rather weak and unrealistic in respect to analytical results, making empirical-computational studies relevant. Further, I found that there are relatively few empirical-computational studies that are in line with the specifications of the uniform taxation theorems mentioned earlier, possibly because of the complexity and heavy informational requirements for computing optimal taxes. But it is also possible that some of the prominent theoretical results in favor of uniformity have been given too much weight, thus retarding further research on this subject. Unfortunately, most empirical-computational models operate under rather restrictive and simplistic assumptions. Although the empirical evidence seems thin so far, it appears that in models that approximate real world conditions, usually a differentiated and progressive indirect tax structure will prevail.10 This is in contrast to the aforementioned policy related studies, which recommend uniform indirect taxation for distributional purposes. The unresolved controversy concerning tax uniformity calls for more empirical-computational research, closer to real world complexity, to find out how the optimal indirect tax structure should look like.

The chapter is structured as follows. Section 2.2 presents the main analytical arguments for and against uniform taxation. Section 2.3 summarizes relevant results from empirical-computational models. Section 2.4 examines the treatment of income tax in optimal mixed taxation models. Section 2.5 takes a critical look at the arguments in favor of uniform taxation that appeared in some recent policy related studies. Some brief conclusions are presented in section 2.6.

---

9 General equilibrium models could be useful when examining distortions on the production side in non-VAT indirect taxation systems. For a theoretical analysis refer to Diamond and Mirrlees (1971).

10 Progressive indirect taxation means higher taxes on luxuries and lower taxes or subsidies on necessities.
2.2 Theoretical considerations

2.2.1 The Ramsey model and its extensions

The original Ramsey model considers a representative consumer economy, where the consumer can allocate his/her total budget between leisure and a number of goods (Ramsey, 1927). The model assumes constant producer prices and no profits. It should be clear from this setup that this model only focuses on the efficiency aspects of commodity taxation. Income tax is assumed to be zero, which is the same as saying that a proportional income tax is allowed. The specifications of the Ramsey model suggest that in this second-best model the objective is to minimize total dead-weight losses (i.e. Harberger triangles). 11

The well-known result from the Ramsey model states the following: At the optimum, a small intensification of all indirect taxes should decrease the compensated demand by the same proportion for all goods. Since substitution effects are associated with efficiency losses, it is not surprising that the focus is on compensated demand. Notice that the Ramsey model does not imply uniform commodity taxation, but only a uniform reduction in compensated demands. 12 In this model uniform taxation will occur only if all compensated elasticities are the same and utility is weakly separable between commodities and leisure.

Diamond (1975) extended the Ramsey model to a many-person model with an endogenously determined redistributive lump-sum grant (b in eq. 1.2), and showed the need for deviating from the principle of 'equal proportional reduction in compensated demand'. The proportional reduction should be less if the consumption of the good concerned is concentrated among groups, i) having a high welfare valuation of marginal income, or ii) having a high propensity to pay taxes. The first condition implies that given two goods with equal compensated demand elasticities, the one consumed more by higher income earners should be reduced more, implying a higher tax rate. The positive correlation between tax rates and consumer group income implies a progressive tax structure. The second condition is closely linked to the shape of Engel curves. Linear Engel curves mean that marginal budget shares are equal, and the propensity to pay taxes will then be equal.

The Ramsey rule only characterises properties of the optimum, and does not provide clear guidance about the optimal tax structure. To examine how the actual tax structure will look like, further assumptions are needed. Sadka (1977) proves that in a single-consumer model, a necessary and sufficient condition for uniformity is that compensated elasticities in respect to the wage rate of different commodities are all equal. This means that a decrease in the wage rate following a proportional income tax increase, will reduce compensated demands by the same proportion. Since this is what characterizes marginal changes at the optimum, proportional income tax represents the optimal solution. A preference structure satisfying this condition is when direct utility can be written as \( U(v(x), L) \), where \( x \) represents commodities; \( L \) is labour supply and \( v(x) \) is a homogenous function. 13 This type of utility function implies Engel

---

11 The first-best solution would be to impose a lump-sum tax.
12 This is also reflected in the numerical results discussed in section 2.3.2.
13 Note, the definition of weakly separable utility is also \( U = U(v(x), L) \), but without a homogeneity condition on \( v \).
elasticities equal to one for all goods.\footnote{Atkinson and Stiglitz (1972) also showed that with weakly separable utility and homogeneous preferences, optimal commodity tax rates will be uniform in a single-person economy. In fact, this theorem applies also in a many-person setting, because given homothetic preferences for goods, differentiated indirect taxation cannot be used as a screening device for ability-to-pay.} It is possible that for some particular groups of goods, a homogeneous sub-utility function will apply. Then these goods should be subject to the same tax rate. But based on empirical evidence, such an assumption cannot be applied to a complete demand system.

Elaborating more on tax structures, Deaton (1981) shows that quasi-separability will lead to a uniform structure in the one-consumer case.\footnote{Quasi-separable utility defines weak separability in respect to the minimum cost function. In this case, the cost function is defined as: $c(u, w, p) = c(u, w, \Omega(u, p))$ where $p$ represents prices, $w$ is the wage rate, $u$ is utility and $\Omega(u, p)$ is a homogeneous function of degree one with respect to prices, representing a perfect consumer price index. Quasi-separable utility is the dual counterpart to weakly-separable utility in the indirect utility space.} Besley and Jewitt (1995) generalize the one-consumer result further and show that it applies to a particular type of utility function. Deaton (1981) also shows that if we move to a many-consumer economy and assume that the planner has preferences in favor of equity, then quasi-separability leads to a regressive tax structure. Weak separability between commodities and leisure leads to a regressive indirect tax structure in the one-consumer case. Introducing an egalitarian planner and many-consumer economy, will move the solution towards progressivity.

Ramsey (1927) and Baumol and Bradford (1970) discuss the so-called inverse elasticity rule for a single-consumer economy, based on the assumption that there are no cross-price effects between commodities, i.e. the demand for all goods depends only on its own price and the price of leisure, namely the wage rate. The rule states that the tax rate should be inversely related to the own price elasticity of a commodity, and will be smaller the more complementary is the commodity with labour. Given that necessities typically have low elasticities of demand, this rule calls for a regressive commodity tax structure in a single-consumer economy. Without cross-price effects, a uniform tax structure will prevail, if all own price elasticities are the same and utility is weakly separable between commodities and leisure.

In a single-consumer and three-good setting (leisure and two commodities), Corlett and Hague (1953) examine which commodities should be taxed to supplement an existing income tax. Their analysis relies on a marginal reform approach, which considers the welfare change due to the introduction of a differentiated commodity tax structure, when the starting point is a uniform system, alternatively interpreted as an existing proportional income tax. They find that the commodity which is a stronger complement to leisure (in Hicks sense, meaning that the compensated cross derivative of the good with respect to the wage rate is negative) should bear a higher tax. This is a marginal analysis, however, not a global one (see section 2.3.6). Still, the result also holds true within an optimal design framework (see Sandmo (1976), Jacobs and Boadway (2014)).

### 2.2.2 The uniform taxation theorems

#### 2.2.2.1 Optimal non-linear income tax and the Atkinson-Stiglitz theorem

A prominent result in the literature is the Atkinson-Stiglitz (A-S) theorem (Atkinson
and Stiglitz, 1976). Conditional on an optimal non-linear income tax, they study the role of indirect taxation in a many-person redistributive model. The starting point for their analysis is the Mirrlees (1971) control theoretic optimal non-linear income tax model, where a uniform lump-sum grant is the main redistributive instrument. Incorporating into the Mirrlees (1971) model non-linear commodity taxes, A-S prove, using control theory, that in this situation optimal non-linear commodity taxes will be all zero, provided preferences are weakly separable between commodities and leisure. This implies that optimal redistribution could be achieved by income tax alone. Given that in the absence of evasion and administration, a portion of income tax can be converted into uniform commodity taxes, the A-S result implies the optimality of zero or uniform commodity taxation.

A particular shortcoming of the original A-S theorem is that it assumes non-linear commodity tax functions, in the form $t_i(x_i)$, where the tax rate $t_i$ is a function of the quantity consumed, $x_i$. This assumption was adopted for reasons of mathematical convenience, because this way the zero commodity taxation theorem could be proved with control theory in a relatively simple manner. But in practice, quantity dependent taxes can be applied only to a few items. They cannot be applied for the majority of goods, because the government cannot properly observe the quantity purchased and consumed by individuals or households, and much of the tax could be avoided through re-trading between consumers. The non-linear commodity tax problem has been corrected by Christiansen (1984), who extends the A-S theorem to proportional commodity taxes (i.e. constant tax rates), by using conditional Marshallian demand functions within a marginal reform framework, in much the same manner as Corlett and Hague (1953). He shows how leisure complements and substitutes should be taxed or subsidised in the presence of an optimal non-linear income tax. He finds that given weakly separable utility and optimal income tax, optimal commodity tax rates should be zero or uniform.

2.2.2.2 The main objections raised against the Atkinson-Stiglitz theorem

A number of objections were raised against the A-S theorem, and here we shall examine them one by one.

Identical preferences for all taxpayers

One of the central assumptions of the Mirrlees (1971) model and the A-S theorem is identical preferences and differing wage rates. However, casual empiricism suggests that taxpayer populations are quite heterogeneous, because of different household compositions and different needs and endowments. The earliest qualification to identical preferences is presented

---

16 A mathematical treatment of the A-S model is presented by Ruiz del Portal (2012) for the non-linear commodity tax case. He introduces into the literature second order effects, including bunching of consumers and kinks in the tax schedule, and a sub-utility function defined as $v(x_1, \ldots, x_n, z)$, which includes income ($z$), in addition to commodities. He finds that the A-S theorem remains valid subject to these complications. Hellwig (2010) presents a similar mathematical analysis for linear taxes.

17 Some mathematical complications that arise due to the application of non-linear commodity taxes in the A-S model are discussed in Revesz (1986).
in Mirrlees (1976) section 4, where he shows that when the population is characterized not by a single characteristic (ability) but by multiple characteristics, then the solution is more complicated, and the A-S theorem will not necessarily apply. Mirrlees applied in this analysis advanced mathematical techniques. Later critiques used simpler mathematics to prove the same point. Saez (2002a) shows, using Christiansen (1984) conditional Marshallian demand functions approach, that the A-S result is no longer valid when considering the heterogeneity of household preferences, reflected by different purchases at the same income level. Other heterogeneity conditions that were shown to invalidate the A-S theorem (using Stiglitz’s “self-selection” approach), include different unobservable endowments apart from different abilities (Cremer et al. (2001)), and different needs, endowments and multiple forms of labour supply (Boadway and Pestieau (2003)). It should be noted that almost all theoretical studies assume a population with identical preferences. Heterogeneous households, in terms of demographic composition, enter into the picture in a few empirical-computational models.

**Weak separability between commodities and leisure**

This is another central assumption in the A-S theorem. Weak separability between commodities and leisure has been rejected in several econometric studies (e.g. Blundell and Walker, 1982; Blundell and Ray, 1984; Browning and Meghir, 1991), thereby invalidating the A-S theorem.

Some recent studies are more ambivalent in relation to commodity/leisure non-separability than those published decades ago. In their contribution to the Mirrlees et al. (2011) Review in relation to VAT and excises (chapter 4), Crawford, Keen and Smith empirically reject weak separability between commodities and leisure, but argue that it is far from clear how much differentiation could be justified on these grounds, or which commodities should be taxed more or less heavily. Pirttilä and Suoniemi (2014) note, that despite the large theoretical literature, there has been little empirical work done in trying to establish the relationship between commodity demand and labour supply. Their research indicates that capital income and housing expenses are negatively associated with working hours, whereas child care is positively related. Generally, it appears that more empirical research is needed on this subject.

**Political acceptability**

In common with many other theoretical models, the A-S and Mirrlees (1971) models assume perfect competition. Yet, in optimal taxation models this assumption has led to rather debatable results. The assumption that wages exactly match productivities, has led in the Mirrlees (1971) model to decreasing marginal income tax rates at the higher income range. Evidently, such a tax schedule is not politically acceptable in a world of imperfect competition and information. This suggests that the Mirrleesian income tax function, which is part of the A-S model, perhaps cannot be implemented in practice (Boadway and Pestieau (2003)).

---

18 Apart from the heterogeneity issue, in section 5 dealing with mixed taxation involving proportional and non-linear commodity taxes, Mirrlees (1976) concluded that if at least one good is subject to a non-linear tax, then proportional taxes should bear more heavily on goods that high ability individuals have relatively stronger taste for. This objection to uniformity is derived from a slightly different model than the A-S theorem.
Actually the situation is more complicated than that. There are many possible shapes of income tax functions that can be derived from the Mirrlees model, depending on the specifications chosen in respect to utility, the distribution of earnings and the distribution of the compensated elasticity of labour supply (see appendix 2.1). Tuomala (2016) examines a number of solutions that appear to be more politically acceptable. These solutions are based on unbounded earning distributions, with an asymptotic upper tail. These solutions show nearly constant or slightly increasing marginal tax rates over most of the income domain. But even in these cases, a Mirrlees type solution cannot be fully implemented in practice. Actual non-linear income tax functions are piecewise linear, rather than smooth differentiable functions of income (Apps et al. 2014). To what extent the uniform taxation theorem will hold true with optimal continuous piecewise-linear income tax functions, is something that has not yet been numerically tested (see appendix 2.2).

2.2.2.3 Non-distributional factors

It is well known from the theoretical literature that differentiated commodity taxation can be justified on grounds other than raising tax revenue for redistribution, such as externalities, (de)merit goods, risks, administrative and compliance costs and tax evasion.\(^{19}\) Sandmo (1975) extends the Ramsey framework to include externalities, and shows that additional taxes equal to the marginal external effects should be added to the original solution. Pirttilä and Tuomala (1997) show how externalities should be taken care of in the presence of a non-linear income tax. Studies taking into account the effects of tax evasion on optimal commodity tax rates include Cremer and Gahvari (1993), Boadway et al. (1994), Ray (1997), Kleven et al. (2000) and myself.\(^{20}\) We shall return to this issue in section 2.5. Cremer and Gahvari (1995) examine the effect of income risks on the optimal taxation of housing and durables. Alm (1996) considers the effect of high administrative and compliance costs on optimal tax rates. Besley (1988), Kanbur et al. (2006) and Pestieau et al. (2012) deal with the merit goods arguments (i.e. paternalistic concerns). Another possible complicating factor is cross-border

\(^{19}\) The Mirrlees model may have had some influence in reducing the progressivity of income tax rates in a number of countries in recent decades, yet, to the best of my knowledge, an income tax schedule based on the Mirrlees model has not been implemented anywhere. Major impediments to its practical implementation include:

- legal-definitional problems related to labour income and to income more generally,
- heterogeneous characteristics and preferences of taxpayers,
- income tax tends to be more vulnerable to tax evasion than most indirect taxes,
- political preferences in favour of using exemptions and concessions with income tax.

\(^{20}\) I present in section 3.5 approximate formulas for the effect of these factors on optimal tax rates.

\(^{21}\) Another interesting tax model on the evasion of small enterprises is presented by Piggott and Whalley (1998). This paper is not included in the list above, because it follows a marginal reform approach, rather than an optimal taxation approach.
shopping. This is examined in Christiansen (1994), Kessing and Koldert (2011) and Nygård (2014). Note that all these complicating factors violate the necessary conditions for tax uniformity.

2.2.2.4 Other possible objections to tax uniformity

There are also other areas where the A-S model is at odds with what is observed in practice. In developed countries, one may observe in-kind transfers (education, health, child care etc.), accessible at no cost or little cost, and different forms of welfare benefits. There are also all kinds of indirect taxes such as VAT, general sales tax, property taxes and excises on specific goods (e.g. gasoline, tobacco, alcohol and motor vehicles). Moreover, governments correct externalities by implementing Pigovian taxes and subsidies. We shall not examine here the optimality or otherwise of different forms of indirect taxation, but a few comments can be made on welfare benefits and in-kind transfers; because this subject has been taken up in some taxation models that follow the Ramsey tradition (see sections 2.2.2.7, 2.3.3.1 and 2.3.3.2).

In common with Mirrlees (1971), the A-S model assumes a single form of income support carried out through an optimal uniform lump-sum grant (called the demogrant). In practice, redistributive support is much more complicated than that. It is made up of cash support through welfare benefits and various forms of in-kind transfers. Piketty and Saez (2013) section 2, note that in developed countries over 50 percent of social programs represent in-kind transfers, rather than cash payments through pensions, child benefits, unemployment benefits and the like. The share of in-kind transfers (mainly education and health services) in social programs is much larger in developing countries, where direct support payments are almost non-existent. The heterogeneity of direct welfare payments can be explained by the heterogeneity of recipient households, a subject that we have already discussed earlier. Given imperfect information by support agencies and possible false reporting by many recipients (particularly in relation to unemployment, disability and means-tested benefits), the perfect optimality of cash payments can be in doubt. The full optimality of in-kind transfers is even more dubious, because here optimality must cover not only scale but also composition.

The implication of sub-optimal lump-sum grants in a heterogeneous population model has been examined theoretically by Deaton and Stern (1986) (see section 2.2.2.7). In the framework of a model based on linear income tax and linear Engel curves, they reach the conclusion that if differentiated lump-sum grants are not provided in optimal scale and composition, then optimal commodity taxation will be differentiated. Moreover, some empirical-computational studies, reviewed in sections 2.3.3.1 and 2.3.3.2, suggest that if the demogrant is set at a sub-optimal level, then the optimal solution will involve differentiated and progressive commodity taxation. While these results do not strictly pertain to the A-S model, they do suggest that some of the many missing elements from this highly abstract model can lead to differentiated taxation. This conjecture should be explored further.

2.2.2.5 The self-selection approach

The A-S theorem can be proved also using Stiglitz’s (1982) “self-selection” optimisation model. This asymmetric information model is applied in a two-person or small group setting. It is a simplified form of the original A-S model. In this type of models, proportional commodity taxes are combined with a non-linear income tax. Marginal income tax rates and corresponding lump-sum taxes or subsidies, are determined at separate income tax
brackets covering individual taxpayers or taxpayer groups. Income and commodity tax rates are chosen so as to ensure that high ability persons will not have the incentive to mimic the incomes of lower ability persons, by reducing their labour supply. The self-selection approach was used in a number of studies that investigated departures from the A-S solution due to various complications. These include Broadway et al. (1994) on the effect of income tax evasion, Cremer et al. (2001) on unobservable endowments in addition to different abilities, Broadway and Pestieau (2003) on different needs, endowments and multiple forms of labour supply, Cremer and Gahvari (1995) on income risks and the purchase of housing and durables, Pirtilä and Tuomala (1997) on externalities and Bastani et al. (2014) on subsidies for child care. Edwards et al. (1994) and Nava et al. (1996) examine the effect of leisure substitution and complementarity on optimal tax rates, as well as their effect on the optimal provision of public goods.

The popularity of the self-selection approach, for the purpose of extending the A-S model in various ways, is motivated by its assumed better tractability compared with the original control theoretic A-S model. While this may well be true in respect to analytical studies, at least for the purpose of numerical studies there seems to be a more direct approach. There is an explicit elasticities based formula for the Mirrlees (1971) income tax problem (see Saez (2001) and Revesz (1989, 2003)). This formula changes slightly following the incorporation of proportional commodity taxes into the income tax model. Denote the marginal commodity tax rate as \( t_{m}^{c} = \sum_{j} t_{j} p_{j} \frac{\partial q_{j}}{\partial m} \) where \( m \) is income, \( q_{j} \) are commodities and \( p_{j} \) producer prices. Then the left hand side of the formula for the optimal marginal income tax \( (T_{m}) \) will be: \( \frac{T_{m} + t_{m}^{c}}{1 - (T_{m} + t_{m}^{c})} \) instead of the original \( \frac{T_{m}}{1 - T_{m}} \) (for proof see appendix 2.1). The right hand side remains the same. The end point conditions at the upper and lower boundaries of the wage distribution of a bounded population will be then: \( T_{m} + t_{m}^{c} = 0 \). Provided \( t_{m}^{c} \) is positive, this implies that at the end points the marginal income tax rate will be negative rather than zero. This result was originally discovered by Cooter (1978). Using the revised optimal income tax formula, it might not be too difficult to investigate various extensions to the A-S model through computational studies. A more robust computational approach is to use in the model a piecewise linear income tax function, instead of the control theoretic method. This subject is discussed in more detail in appendix 2.2.

2.2.2.6 The Laroque-Kaplow proposition: Sub-optimal income tax

The Laroque-Kaplow (L-K) proposition extends the A-S theorem to apply to non-optimal income tax functions as well (Laroque, 2005; Kaplow, 2006). It states that with weak separability and identical preferences, it is possible to replace any non-uniform indirect tax structure by zero or uniform commodity taxes, by adjusting the non-optimal income tax function and the demogrant in such a way that all taxpayers will maintain or improve their utility position. Hence, the L-K proposition suggests that in a redistributive model, eliminating

---

22 The income tax functions obtained from the self-selection models are piecewise-linear, however, they differ from actual income tax schedules that are also piecewise-linear. Actual income tax schedules are continuous functions of income, while those obtained from the self-selection model may be discontinuous (see appendix 2.2).
non-uniform commodity taxes can lead to a Pareto improvement. The L-K proposition shares all the empirical objections to the A-S theorem outlined in sections 2.2.2.2 to 2.2.2.4. But it has some other weaknesses as well.

Boadway (2010) examines this proposition and stresses that if because of some reason (political or administrative), the appropriate income tax adjustment is not carried out, then such a reform may be welfare reducing. Since the L-K proposition does not describe the design of the income tax functions that combined with zero or uniform commodity taxes will yield improved welfare, this critique seems pertinent. The discussion in section 3.4.5 suggests that the mathematical theorems presented by Laroque (2005) and Kaplow (2006), do not really prove that following the L-K reform a Pareto improvement will necessarily occur. The problem seems to be that Laroque and Kaplow did not recognize properly that the work related disincentive effects of direct and indirect taxes are substantially different.

2.2.2.7 Optimal linear income tax and the Deaton theorem

Assuming that income tax is restricted to be linear, Deaton (1979a) proved that in a many-person model with identical preferences, weak separability and linear Engel curves, the optimal commodity tax structure will be uniform. This result is similar to the A-S theorem, but given that income tax is linear, the assumption of linear Engel curves must be added in order to obtain a uniform commodity tax solution. The assumption of linear Engel curves for all goods is dubious and has been rejected in econometric studies; see e.g. Blundell and Ray (1984). Revesz (1997) shows that in a multi-product setting, because of the non-negativity constraint on demand, linear Engel curves for all goods imply nearly homothetic preferences, which is clearly unrealistic. The Deaton (1979a) model shares all the empirical objections to the A-S model outlined in sections 2.2.2.2 to 2.2.2.4, apart from the difficulties in implementing an optimal Mirrleesian income tax schedule. However, an extension of Deaton’s theorem takes care of heterogeneous populations, as will be explained below.

Deaton and Stern (1986) assume weakly separable utility, linear Engel curves and linear income tax, and that consumers differ in preferences (and consumption patterns) partly due to differences in observable policy related characteristics (such as age or number of children) and partly due to idiosyncratic preference variation. Differences in preferences are represented in their model by differences in the intercepts of the Engel curves. They show that if i) social valuation weights are correlated with differences in preferences and characteristics, ii) variations in consumption are related to differences in policy related characteristics, and iii) lump-sum grants can be conditioned on policy related characteristics in a linear way, then uniform indirect taxation is optimal. In this case, we only need to design an optimal set of lump-sum grants dependent on policy related characteristics and differentiated indirect taxation is superfluous. Let me just note here that due to imperfect information, the possibility of optimal

23 Looking from a different perspective, Revesz (1997) demonstrates analytically and with numerical results that a revenue-neutral replacement of progressive indirect taxation by progressive direct taxation, which leads to zero commodity taxes, will reduce labour supply and social welfare. This finding also appears to contradict the L-K proposition.

24 In the Deaton (1979a) model, in order to have the same propensity to pay taxes at all income levels, (i.e. have constant marginal indirect tax rate), all linear Engel curves must start with positive values from the lowest disposable income.
lump-sum grants is somewhat problematic from a real world perspective, and as indicated earlier, the assumption of linear Engel curves for all goods is not supported by empirical evidence. The Deaton-Stern (1986) theorem has been supported by the numerical results of Ebrahimi and Heady (1988) (see sections 2.3.3.2 and 2.3.3.3).

In a recent paper, Boadway and Song (2016) examine analytically certain extensions to the Deaton (1979a) model. They investigate a two-good model, where one good is a necessity and the other is a luxury. They find that if income tax (linear or non-linear) is less progressive than optimal, then the necessity should be taxed at a lower rate than the luxury. Given that a commodity tax model where only a lump-sum grant is present but no income tax, is effectively a model with a particular form of linear income tax (see section 2.3.1), the conclusion of Boadway and Song applies to these models as well. In section 2.3.3.2 we shall examine empirical-computational models of this type, where the lump-sum grant is sub-optimal. The numerical results from these models accord with the Boadway-Song theorem.

Another conclusion presented in their paper is that if a linear income tax function is optimal, but low-income households are unable to afford luxury goods, it may be optimal to tax necessity goods at lower rates than luxuries. My computational studies, discussed in chapter 3, examined the case of non-linear Engel curves, where luxuries are consumed only by high income earners. The numerical results from these models (summarised in section 2.3.3.3) lead to the same conclusion as Boadway and Song (2016).

2.3. Empirical-computational studies

2.3.1 Methodological issues

To apply the theory in empirical-computational models raises several questions: What kind of information is needed? Which type of utility function should be used? What kind of methods and specifications should be applied? The first order conditions and the budget constraint will make our need for information about the individuals' demand and its derivatives immediately apparent.

We will also need information on individuals' supply functions, i.e. labour supply. Without imposing further restrictions, such as commodity/leisure separability, the wage rate and leisure consumption will influence the demand for commodities through substitution and complementarity and not only income effects. Given that separability between commodities and leisure has been rejected in several econometric demand studies (see section 2.2.2.2), it does not appear to be a realistic restriction. On the other hand, there are few reliable estimates on leisure substitution or complementarity parameters (Jacobs and Boadway, 2014).

In line with earlier discussion, one could argue that the functional form of demand should be flexible enough to allow for non-linear Engel curves and non-separable utility. Yet we should be aware that undertaking an optimal design analysis requires the demand and labour supply functions to be consistent with consumer theory globally and not only locally, since the optimal price structure could be far from the point at which the functions are consistent with theory. So-called flexible functions do not automatically exhibit these properties in a global
sense. The property of a flexible functional form is its ability to take on any set of price and income elasticities at a particular data point, unrestricted by a priori assumptions. This seems very desirable, but it comes at a cost. As Caves and Christensen (1980, p. 423) make it clear, outside the initial data points the estimated flexible form indirect utility function may not be monotonic or strictly quasi-convex (implying quasi-concavity of direct utility). An example of a popular and widely used flexible functional form is the almost ideal demand system (AIDS) introduced by Deaton and Muellbauer (1980). This function may provide an approximation to any arbitrary utility function and may exhibit many desirable properties locally, such as the quasi-convexity of indirect utility. However, it can only show consistency with demand theory locally, and there is no guarantee for the same applying globally.

In addition, in a redistributive model the social welfare function must be specified in detail for a many-consumer economy. This raises issues in regard to the cardinalisation of utility and political value judgments, which are usually captured by the inequality aversion rate. Household composition also raises inter-personal comparability issues. One way to resolve the problem is to assign politically determined utility weights according to household composition. A more objective approach is to use demographic equivalence scales, derived from household expenditure surveys (Ray, 1989). Determining the equivalence scales requires the decomposition of average household expenditure between household members, divided into categories such as parents, non-parents, children by age group, etc. In addition, we have to estimate economies of scale in consumption, depending on household size. Whether demographic equivalence scales can be determined objectively, is an open question.

Most of the numerical models on commodity taxation published so far exclude income tax, or at least the variable part of income tax. Given a linear income tax function commonly defined as: \( \alpha + \beta m_h \), then \( \beta m_h \) represents the variable part. Models where only \( \alpha \) is present are in effect also linear income tax models. For example, assume an initial situation where income tax is \( \alpha + \beta m_h \) and all commodity taxes are zero. Then a transformation reduces the variable income tax rate (\( \beta \)) to zero, while at the same time it increases commodity taxes (relative to producer prices) and the lump-sum grant by the factor \( 1/(1 - \beta) \). This will effectively not change anything, by virtue of zero homogeneous utility and demand. But in the new situation, only the lump-sum grant, \( \alpha/(1 - \beta) \), is left in the model, which will then represent a particular form of linear income tax.

After a utility function has been selected and all the information needed is at hand, we can calculate optimal taxes. This is not a trivial task since taxes, prices, quantities and elasticities are interdependent in a non-linear way. Some kind of iterative numerical method must be employed to yield a solution. Given quasi-concave preferences and a convex budget set, a `gradient` based iterative approximation will lead to the globally optimal solution (see

---

25 Utility functions with few parameters, such as LES, CES and Cobb-Douglas, satisfy global compatibility with demand theory (including global quasi-concavity) with ordinary parameter estimates.

26 AIDS utility is defined in terms of the minimum cost function as:

\[
\log m = \alpha_0 + \sum_i \alpha_i \log p_i + \frac{1}{2} \sum_{ij} \gamma_{ij} \log p_i \log p_j + \sum_i \beta_i \prod_j p_j^{\beta_i} \quad s.t. \quad \sum_i \beta_i = 0; \quad \sum_i \gamma_{ij} = 0; \quad \gamma_{ij} = \gamma_{ji}
\]

with corresponding demand: \( w_i = \alpha_i + \sum_j \gamma_{ij} \log p_j + \beta_i \log (\sum_p m) \)

where \( m \) is total expenditure, \( w \) is budget share, \( U \) is utility, \( p_i \) are prices and \( P \) is a general price index.
appendix 2.2 and section 3.3). If we relax the assumption of constant producer prices, the computation of optimal taxes will also require information about the producer sector, which would render the task even harder.

Having reviewed some of the difficulties with empirical-computational studies, we can now turn to examine a number of contributions. To my knowledge, apart from the self-selection model of Bastani et al. (2014), no attempt has been made to test numerically the theory of non-linear income tax along with commodity taxation according to the ‘Mirrlees tradition’. We shall therefore review mainly studies under the extended Ramsey model, i.e., linear or proportional taxation. The studies reviewed consider only the demand side, assuming constant producer prices. With the exception of four studies, the demographic composition of households is excluded from these models.

2.3.2 Representative consumer models

We start the review of empirical-computational models with a number of studies that are considering a representative consumer economy. To what extent these one-person models are relevant to distributional models is an open question. Whatever the case may be, they illustrate various statistical and mathematical approaches to the commodity tax optimisation problem.27

Atkinson and Stiglitz’s (1972) paper is the first, to my knowledge, to compute optimal tax rates from empirical data. In computing optimal taxes for five commodity groups they consider two demand systems: The linear expenditure system (LES) based on estimates by Stone (1954), and the direct addilog demand system based on estimates by Houthakker (1960).28 In both cases there is separability between commodities and leisure and they assume elastic labor supply. According to theory they should get a solution which is regressive, and so they do.29

Fukushima and Hatta (1989), using the same data set and the same model, find that reducing the (compensated) labour supply elasticities works in favor of a uniform system. With what they consider as more reasonable values, they find the structure to be fairly uniform.

Harris and McKinnon (1979) also calculated optimal tax rates for five product groups, using a Stone-Geary (LES) function with leisure and commodities. The optimal structure, they conclude, varies with the assumed compensated labour supply elasticities. Fukushima (1991) uses the same data but with lower labour supply elasticity, which yields a result somewhat closer to a uniform solution.

---

27 In terms of analytical structure, these models are close to Ramsey (1927). In all these models there is a given tax revenue requirement.

28 The Stone-Geary linear expenditure system (LES) is defined as:

\[
\text{utility} \quad u = \sum \beta_i \log(x_i - y_i) \quad \text{s.t.} \quad \sum \beta_i = 1
\]

29 For these theoretical results refer to Deaton (1981) and Baumol and Bradford (1970).
Asano and Fukushima (2006) estimate the joint decision of leisure and commodity demand without imposing any separability restriction. They use Deaton’s AIDS and compute optimal tax rates for ten commodity groups in Japan. Their conclusion is that the optimal structure is reasonably close to a uniform one, which suggests that the welfare losses associated with tax uniformity are small. Although the Asano and Fukushima model provides important improvements, their results are weakened by uncertainty to what extent the model complies with properties required according to demand theory (see section 2.3.5).

Another study within the one-person economy framework is Nygård (2014). Using a LES system, he includes cross-border shopping and focuses on cross-border exposed goods for Norway. The goods purchased across the border are assumed to be non-taxable and externality-generating. He optimizes commodity taxes conditional on a pre-existing income tax. As expected, he shows that goods purchased at home should be taxed more leniently, because of the distortions caused by cross-border shopping. In particular, he shows how the effects get stronger because these goods are externality-generating. In general, the tax structure obtained is highly differentiated. When neglecting cross-border shopping and external effects, he gets a more uniform solution, though still more differentiated than that of Asano and Fukushima (2006).

2.3.3. Many-consumer economy

2.3.3.1. Models without a lump-sum grant

In almost all theoretical models, redistributive support is represented by a single uniform lump-sum grant for everyone, called the demogrant. In a number of empirical-computational models the demogrant is missing. This is supposed to represent the situation in many developing countries where direct support payments are absent. Looking back at eq. (1.2), the budget constraint in these models appears as: $\sum_{i}^{N} \sum_{h}^{H} t_{i} p_{i} x_{ih} = 0$. Hb is missing because b is set to zero. $T(m_{h})$ and G are missing because in the models described below, there is no income tax or fixed public goods expenditure requirements. Given the above constraint, commodity taxes and subsidies must add up to the same totals. The presence of egalitarian objectives and the absence of a demogrant, imply progressive indirect taxation/subsidisation at the optimum, because there is no other way to provide support to the needy, as limited as it may be. Here I shall review some of these models.

As far as I am aware, the first calculation of optimal taxes within a many-person framework is presented in Deaton (1977). Deaton's model relies heavily on simplifying assumptions. By employing what he calls strategic aggregation, he ends up having to consider the behavior of only two consumers, the marginal utilities weighted social representative consumer and the average consumer. He calculates optimal taxes for eight goods. His study is based on inelastic labour supply and linear Engel curves. His specification of the welfare function is based on Atkinson (1970), and is similar to the welfare functions used in most of the studies that will be reviewed later. In the absence of distributional goals, his results indicate a uniform tax structure. He finds that when the concern for equity increases, the structure becomes more differentiated and luxuries are taxed more heavily than necessities.

Heady and Mitra (1980) use a Stone-Geary LES utility function, implying both separability and linear Engel curves for nine goods, including leisure. Basically they find that
the tax structure is progressive, depending on what assumption is adopted about equity.

Sensitivity of optimal tax rates to different demand systems is considered by Ray (1986), in a model that excludes leisure. He calculates optimal tax rates for nine goods from Indian data, conditional on the prices, incomes and elasticities observed at a particular point of time. That means that his optimal tax rates are not optimal in a strict sense, but only reflect what the tax rates would have been had the initial situation constituted the optimum.\(^3\) He compares the linear expenditure system (LES) with the restricted non-linear preference system (RNLPS), which is a specialisation of the non-linear preference system (NLPS) introduced by Blundell and Ray (1984).\(^3\) The RNLPS allows for non-linear Engel curves. He finds that results from the two demand systems agree at low level of concern for equity, but diverge when the inequality aversion rate increases. At low levels of inequality aversion they approach a uniform solution. Ray (1986) also finds that the scale of redistribution achieved through the taxation/subsidisation of nine broad product groups is fairly limited.

A study that follows up Ray (1986) is Majumder (1988). Using the same conditional method, he tests other non-linear Engel curves specifications on the same data, and discovers that the results are sensitive to the exact specification of Engel curves non-linearity.

Murty and Ray (1987) use the general functional form of Blundell and Ray (1984), the non-linear preference system (NLPS), to investigate sensitivity to the assumption of weakly separable utility between goods and leisure. Their results indicate that the tax rates for the nine goods considered are highly sensitive to deviations from weak separability.

Murty and Ray (1989) develop an iterative algorithm, based on the marginal tax reform approach of Ahmad and Stern (1984), to calculate optimal tax rates. Because of the iterative calculations, their approximation to optimal tax rates is probably better than in some of the studies mentioned earlier.

Ray and Blacklow (2002) extend this model and incorporate demographic effects, when they use RNLPS and LES to study optimal taxes in Australia for nine goods. The optimal tax rates, they conclude, move away from uniformity when demographics are introduced and affect the social welfare weights. The effect is more significant when considering the RNLPS than LES. In line with Murty and Ray’s (1989) findings, their results also indicate that LES and RNLPS agree at low inequality aversion, but diverge at higher inequality aversion rates. Furthermore, optimal tax rates appear to be more sensitive to the choice of functional form than to the inclusion of demographic effects.

It should be noted that a redistributive model without a lump-sum grant is hardly

---

\(^{30}\) It appears that Ray (1986) performed only the first step of what should be a multi-step iterative computational process.

\(^{31}\) The cost function underlying the NLPS is written as \(c(u, p) = \left[ a(p, \alpha) + ub(p, \alpha) \right]^{1/\beta} \), where \(a(.)\) and \(b(.)\) are homogeneous of degree \(\alpha\) in prices, \(p\). When \(\alpha = 1\) this reduces to the Gorman Polar Form family of systems with LES being the most well-known. We obtain the NLPS by specifying

\[
\begin{align*}
\alpha & = \sum_{i=1}^{n} \frac{\alpha}{Y_{ij}} \sum_{j=1}^{n} \beta_{ij} P_i^\alpha P_j^\beta, \\
\beta & = \prod_{k=1}^{n} \beta_k \text{ and } \sum_{k=1}^{n} \beta_k = 1.
\end{align*}
\]

We can then derive the NLPS demand function in budget share terms, \(w_i\), as:

\[
w_i = \sum_{j=1}^{n} \frac{\alpha^{Z_{ij}} Y_{ij} / (2 \beta_{ij} Z_{ij}^{\alpha} / 2 + 1)}{\sum_{j=1}^{n} \frac{\alpha^{Z_{ij}} Y_{ij} / (2 \beta_{ij} Z_{ij}^{\alpha} / 2)}} \text{, s.t. } \sum_{i=1}^{n} \beta_i = 1, Y_{ij} = y_{ij}, \text{ where } Z_i = P_i / x
\]

and \(x\) is aggregate expenditure. The RNLPS follows from this when setting \(y_{ij} = 0, \text{ for } i \neq j\). RNLPS still allows for non-linearity but not for non-separability.
realistic in any country. Even in developing countries where direct support payments are absent, there is some kind of redistributive support provided through public education, health care and other in-kind transfers, which can be represented for modelling purposes by a demogrant. Alternatively, the set of subsidized goods in a zero lump-sum grant model should include education, health, child care and public transport, provided there is some information on their consumption by income groups and on their demand elasticities.

2.3.3.2 Models with sub-optimal lump-sum grants

Another class of empirical-computational models deals with positive but sub-optimal lump-sum grants. The budget constraint of such models with identical preferences appears as: \( \sum_i^{N} \sum_h^{H} t_i p_i x_{i,h} = H b \). But in this case the demogrant (b) is not the solution from full welfare function optimisation. In this section we shall only deal with cases where “b” is below the optimal value, and the variable part of income tax (\( b m_h \)) is missing.

The demogrant can be sub-optimal for two reasons. The first is because it is given exogenously below the optimal level. When the lump-sum grant is sub-optimal, it is reasonable to assume that progressive indirect taxation is needed to compensate for sub-optimal support provided to the needy. Sub-optimal demogrants can represent the situation in countries, where because of the proliferation of the “shadow economy”, and/or because of serious wastage in the public sector, expenditure on redistribution does not reach the politically preferred level (section 3.4.4).\(^{32}\) The theoretical study by Broadway and Song (2016) (discussed in section 2.2.2.7) found that given a sub-optimal linear income tax, optimal commodity tax rates should be progressive. Because of the absence of the variable part of income tax in the models described in this section, the sub-optimal demogrant alone represents in effect a form of sub-optimal linear income tax (see section 2.3.1), and indeed, the numerical results from these models accord with the Broadway-Song theorem.

Another reason why the demogrant can be sub-optimal arises in populations with heterogeneous characteristics. Under these conditions, varying amount of support should be provided, depending on policy related characteristics of the household. In this case the budget constraint will be: \( \sum_i^{N} \sum_h^{H} t_i p_i x_{i,h} = \sum_h b_h \). When the optimisation does not take into account all the information available in the model about household characteristics, then the resulting lump-sum grants will be sub-optimal. For example, in the model of Ebrahimi and Heady (1988), there are two rates of child benefits, depending on the age of the child. When the two rates are reduced to one (the sub-optimal case) then optimal commodity tax rates become differentiated. This result confirms the theoretical contribution of Deaton and Stern (1986).

Ebrahimi and Heady (1988) and Ray (1989) investigate the impact on optimal tax rates of child benefits. Both of them use data from a UK database covering four composite goods. Ray (1989) allows for non-linear Engel curves but does not include leisure, while Ebrahimi and Heady (1988) include leisure and linear Engel curves that are not always parallel across households. Since Ebrahimi and Heady combine a many-person heterogeneous population model with variable labour supply and differentiated lump-sum grants, their study is probably one of the most sophisticated and comprehensive empirical-computational models published so

---

\(^{32}\) Chapter 3 Table D.2 presents some illustrative calculations on the dispersion of commodity tax rates in the presence of sub-optimal lump-sum grants, starting from zero and up to the optimum.
far. We shall return to their study in the next section. Ray (1989) finds support for progressivity when lump-sum grants (made up only of child benefits) are sub-optimal, especially when inequality aversion is high. Ebrahimi and Heady (1988) conclude that commodity tax rates will not be uniform, when the demogrant (lump sum grants conditioned on demographic characteristics of the household plus a uniform component) are not set optimally, or if preferences are not weakly separable between commodities and leisure.\(^{33}\)

Asano et al. (2004) use an AIDS system without variable labour supply, which has much in common with Ray (1989). If the lump-sum grant is zero or is set at a sub-optimal level, they get a progressive structure, in line with Ray (1986, 1989). When the uniform lump sum grant is free to change in the course of tax optimisation, they get a regressive tax structure. Given that in their model labour supply is fixed, it is not clear how they obtained finite tax rates from full optimisation. In a redistributive model where the only primary input (labour) is fixed, one would expect the optimal solution to tax away and redistribute all incomes equally through the variable demogrant (implying infinite commodity tax rates), since output would be unaffected.

### 2.3.3.3 Models with endogenously determined optimal lump-sum grants

In these models the lump-sum grant is free to change in the course of optimisation, in line with the specifications in eq. (1). In most of the scenarios described below, the budget constraint is similar to eq. (1.2), although in some of these models the variable part of income (\(\beta m_h\)) is missing.\(^{34}\) As explained in section 2.3.1, a model where \(\alpha\) is present but \(\beta m_h\) is absent, implies a particular form of linear income tax.

Ebrahimi and Heady (1988) present some scenarios where lump-sum grants are determined through full tax optimisation, taking into account all demographic characteristics and the tax revenue constraint. Under these conditions, they find that optimal commodity tax rates will be uniform, provided utility is weakly separable between commodities and leisure, and provided Engel curves are linear and parallel across households, in line with the analysis of Deaton and Stern (1986). If either of these conditions is violated, optimal tax rates will not be uniform. In their model, non-parallel Engel curves represent a mild case of non-linearity. But even that slight non-linearity of Engel curves has led to perceptible differentiation in tax rates.

Revesz (1997) uses LES first with nine goods for everyone, then with 9 goods for the poor and 18 goods for higher income earners. In both cases leisure is separable from other goods and a uniform demogrant is calculated from the tax revenue constraint. In the 9 goods model I obtained results in line with Deaton (1979a), namely uniformity. When applying the 18 goods setup, I obtained a progressive structure, because of the non-linearity of Engel curves. This model illustrates the commonly observed situation, where the consumption of luxuries starts at intermediate or high income levels. The 9 goods parameter estimates used in this study were

---

33 While the optimal commodity tax rates reported by Ebrahimi and Heady (1988) are usually fairly dispersed, they note that the welfare gains of the optimal non-uniform solution over the uniform one turned out to be quite small, in the order of less than 0.1 percent of total income. This result may be due to the fact that they employed a linear Engel curves demand system. As shown in chapter 3, similar calculations with non-linear Engel curves can lead to much higher welfare gains, sometimes reaching 3 percent or more of total income.

34 These include Ebrahimi and Heady (1988) and Revesz (1997).
Revesz (1997) has been developed further, to become a fairly comprehensive numerical model of optimal indirect taxation in Revesz (2014a, 2014b), as described in chapter 3. Unlike other numerical studies mentioned earlier, I did not calibrate the calculations to empirical data. Instead, I used in my numerical examples arbitrary but plausible numbers. I justified this approach by presenting approximate formulas for optimal commodity tax rates. The numerical results are used to illustrate and substantiate the analytical approximations. As explained in Chapter 3, this model finds that under non-linear Engel curves and logarithmic utility specifications and without income tax, optimal commodity tax rates are highly differentiated and progressive. The dispersion of tax rates is further increased in the presence of “real life” complexities, such as evasion, administrative costs, externalities and non-separable utility between commodities and leisure. The dispersion of tax rates is reduced if an exogenously given non-linear income tax function is incorporated into the model, or if the inequality aversion rate is low, or the average compensated elasticity of labour supply is low.

A computational study by Bastani, Blomquist and Pirttila (2014) examines the effect of strong leisure substitutes, such as child-care and aged-care services on optimal commodity tax rates, using a Stiglitz (1982) type self-selection model. In the self-selection model the zero intercept of the optimal non-linear income tax function represents an endogenously determined lump-sum grant. The model assumes that an hour child care will generate an additional hour of work. Bastani et al. (2014) examine optimal taxes-subsidies and changes in labour supply in a 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.

2.3.4 Relevance to the uniform taxation theorems

Having reviewed the empirical-computational models published so far, the question arises to what extent are these models relevant to the uniform taxation theorems? The models covered in section 2.3.2 (single-consumer economy) and section 2.3.3.1 (the absence of income tax and a lump-sum grant) appear far removed from the specifications in eq. (1) of the A-S and Deaton models. The framework of the sub-optimal demogrant models, discussed in section 2.3.3.2, is already closer to those of the uniform taxation theorems. In fact, there are two theoretical studies – Deaton and Stern (1986) and Boadway and Song (2016) - that are directly relevant to these numerical models. The models in section 2.3.3.3, with endogenously determined optimal lump-sum grants, are perhaps the closest to the framework of the uniform taxation theorems. But even here, some departures from the original specifications can be found. All the models in section 2.3.3.3 are based on populations with non-identical preferences. With the demographic models of Ebrahimi and Heady (1988) and Bastani et al. (2014), preferences are heterogeneous because of different household characteristics. In my models (described in chapter 3), two different sets of preferences are specified for low and high-income taxpayers, in order to obtain non-linear Engel curves with LES. Such a segmented utility framework is not really needed for non-linear Engel curves, but it is necessary in these particular
models. Despite some departures from the specifications in eq. (1), the models discussed in section 2.3.3.3 appear to be quite relevant to the tax uniformity debate.

2.3.5 Some critical remarks on flexible demand functions

The demand systems used in early empirical-computational work (such as LES) relied on very strict assumptions, such as separability and linear Engel curves. This does have the advantage of being perfectly consistent with consumer theory. More recent studies have avoided putting a priori assumptions on behavior, by using flexible functional forms. The results risk inconsistency with consumer theory. While this issue is neglected in the literature, several examples can be given.

The study of Asano and Fukushima (2006) shows important improvements (such as increasing the number of goods and not imposing separability, together with non-linear Engel curves), but their results are weakened by the fact that nothing is done in order to check global characteristics at the optimum. They discuss theory consistency for estimated elasticities at the sample mean, but make no attempt to clarify whether the same conditions will be fulfilled at the optimum.

The same criticism applies to other studies, such as Murty and Ray (1989) and Ray and Blacklow (2002). The flexible form demand system, NLPS, is not globally quasi-concave (Blundell and Ray, 1984, p. 802). How the restricted versions (such as RNLPS) perform in this regard has not been explained.

2.3.6 An alternative approach: Optimal marginal reforms

Another numerical approach to tax optimisation is to consider marginal reforms instead of globally optimal tax designs. Ahmad and Stern (1984) method is based on the marginal utility of public funds, defined for each commodity as: \( \lambda_k = \frac{\partial W}{\partial \tau_k} / \frac{\partial R}{\partial \tau_k} \), where \( W \) is social welfare and \( R \) is total tax revenue. At the global optimum all \( \lambda_k \) will be equal, otherwise they will be different. Raising taxes on goods with low \( \lambda_k \) and reducing them on those with high \( \lambda_k \) will increase social welfare, indicating the advisable direction for tax reform, but not its scale.

From a practical point of view, this approach has the advantage of considerably lower information requirements than full optimisation. Whereas the globally optimal tax approach demands knowledge about individuals’ complete demand system, the marginal reform approach only needs information about individuals’ consumption expenditures, aggregate demand derivatives and tax rates in the initial situation (see Santoro (2007) for a survey). Moreover, the need for global consistency with demand theory is not present, so flexible functional forms could be used with more ease. The marginal reform approach also seems to be more robust to the choice of specifications than the globally optimal tax approach (Madden, 1995; Decoster and Schokkaert, 1990). All this makes marginal reform analysis attractive, but it remains

---

35 Had these models applied identical LES preferences, some luxury goods would show up with negative demand for low income consumers. Setting these negative values to zero would upset the budget constraint. Thus, two separate LES utility functions were needed in order to resolve this problem.
somewhat limited in scope. It only indicates the direction for welfare improvement, without ensuring that the solution is the best possible outcome.\textsuperscript{36}

Marginal reform studies tend to support non-uniform commodity taxation. This review has not examined these studies in detail, because the focus here is on optimal taxation, given that the uniform taxation theorems assume full optimisation.\textsuperscript{37} The marginal reform approach is essentially a comparative static exercise. Yet, it should be noted that the marginal reform methods could have a significant role in indicating the advisable direction for tax reform, under conditions of limited information.

\section{2.4 The treatment of income tax}

A particular problem, that diminishes the practical relevance of both the theoretical and the empirical-computational literature, is the treatment of income tax. In many computational models income tax is missing entirely. Almost all theoretical studies deal with A-S model or its simplified version, the Stiglitz self-selection model. As noted in section 2.2.2.2, certain solutions to the Mirrlees (1971) model may not be politically acceptable. Furthermore actual non-linear income tax schedules are piecewise linear and not smooth differentiable functions, as assumed in the Mirrlees model (see appendices 2.1 and 2.2). The income tax schedules emerging from the Stiglitz self-selection model are even less realistic. For example, in a two-person model, the marginal income tax rate on the higher ability person will be zero and on the lower ability person it will be positive, with a lump-sum tax separating the two.

A linear income tax appears more realistic. It is used in a number of empirical-computational models, apart from the single-consumer and the zero lump-sum grant models. Nonetheless, it still departs from the piecewise linear income tax schedules encountered in practice. Given that income tax has a strong influence on the structure of optimal indirect taxation, the question arises, what would be the structure of indirect taxes in the presence of actual direct tax schedules (broadly defined, including social security contributions)? There are very few papers that touch on this subject. Section 3.4.2 incorporates hypothetical piecewise-linear income tax schedules into some scenarios. The results indicate that the presence of a nonlinear and non-optimal income tax schedule reduces the dispersion and progressivity of optimal commodity tax rates. Given the worldwide trend to reduce and flatten income taxes, this finding suggests that following the reduction of income tax, the progressivity of indirect taxation should have been increased. In view of the practical difficulties in implementing a Mirrleesian optimal non-linear income tax, there seems to be a need to construct numerical models of optimal indirect taxation that incorporate actual or proposed income tax schedules.

Apart from greater realism, there is another advantage in using indirect tax optimisation models with actual income tax schedules, plus a lump-sum grant that represents actual social spending per person (including welfare payments and in-kind transfers), and an actual figure

\textsuperscript{36} The marginal reform method could be also used in an iterative numerical search for the global optimum (see Murty and Ray, 1989).

\textsuperscript{37} Marginal reform can be used to demonstrate the superiority of non-uniform taxes compared to the uniform solution (i.e. Madden, 1995).
for expenditure on public goods.\(^{38}\) This way, the optimal solution for indirect taxes will be, to some extent, free of value judgments. With \(T(m_h)\) and \(b\) fixed in equations (1.1) and (1.2), the problem is no longer joint optimisation of income and commodity taxation, but the optimisation of indirect taxes alone. In this case another constraint is added. Not only must the solution satisfy eq. (1.2), but it must also satisfy the condition that the \(b\) obtained from the solution must equal to the actually observed lump-sum grant per person (denoted \(\hat{b}\)). Assuming that the social welfare function (\(W\)) is characterized by a single inequality aversion rate, the way to ensure that both eq. (1.2) and \(b = \hat{b}\) are satisfied, is to run optimisations with different values of the inequality aversion rate, and then select the solution that yields a lump-grant equal to \(\hat{b}\). This will be the required optimal solution. This solution simply represents optimal indirect tax rates under given empirical conditions. Distributional value judgments are reflected in \(\hat{b}\) and \(T(m_h)\).

### 2.5 Uniformity and policy related studies

In recent years some important policy related studies have supported uniform indirect taxation. These include Mirrlees et al. (2011), Arnold et al. (2011), European Commission (2013), IMF (2014) and NOU (2014). The common rationale is that income taxation and welfare payments are more suitable tools for redistributive purposes than progressive indirect taxation. This reasoning is in line with the A-S, L-K and Deaton theorems mentioned earlier.\(^{39}\) Yet, these analytical results are based on empirically incorrect assumptions, as explained in section 2.2. The well-known labour and saving disincentives and evasion-avoidance problems associated with income tax, and false reporting with welfare payments, cast doubts about their suitability to address fully all distributional objectives. It is paradoxical that while these policy studies (with the possible exception of Mirrlees et al. (2011)) call for a reduction in income tax and its replacement by consumption taxes, at the same time, they expect income tax to be the sole tax instrument to mitigate distributional problems. Some of these policy studies also mention other factors in favor of uniform taxation. These supposed benefits include:

- More effective support to the needy through lower income tax rates or higher welfare payments than through reduced taxes on necessities.
- Lower administrative costs with uniform indirect taxation.
- Narrowing the scope for tax evasion with uniform tax rates.
- In non-VAT systems, due to cascading taxes on inputs, efficiency losses may arise as a result of non-uniform taxes on the inputs to production.\(^{40}\)

---

\(^{38}\) For a more realistic representation, \(b\) can be made a function of income, i.e. \(b(m_h)\). With this setup, \(\hat{b}\) will equal average \(b(m_h)\).

\(^{39}\) Generally, most policy related studies accept that there is a role for Pigovian taxes to correct negative externalities, particularly on the environmental front. However, there is a less supportive view on differentiated indirect taxation, when it comes to distributional or paternalistic concerns, or to leisure complements and substitutes.

\(^{40}\) Examination of this issue requires either general equilibrium modelling, or a comparative analysis of these production efficiency losses with the gains to consumers from a simpler non-VAT sales tax system.
Section 3.5 and appendix 2.2 suggest that computational models can be made as complex as required for realistic representation. In elaborate models, the abovementioned issues could be examined quantitatively through appropriate enhancements. There is no need to relegate these subjects to intuitive discussions separately from other issues. Arguably, policy analysis should rely more on empirical estimates combined with comprehensive computational modelling.

Some policy related studies note that even if sometimes there are reasons to adopt non-uniform commodity tax rates, the gains from such a policy (apart from rectifying externalities), are unlikely to outweigh the cost of a more cumbersome tax administration due to non-uniform taxes. This opinion is echoed in a number of studies, including Mirrlees et al (2011), Piketty and Saez (2013), Keen (2013) and IMF (2014). But here the questions arises, is this view well-founded on empirical observations? Section 3.5.3 points out that there is evidence to suggest that goods and services produced and marketed by large organisations, tend to be less evasion prone than those produced and marketed by the self-employed or small business. Other goods with lower evasion propensities, include products where a large percentage passes through border checkpoints, and highly visible goods, such as real estate and motor vehicles. In view of widely differing evasion propensities, the assumption that uniform tax rates will lead to a less costly and more effective tax administration, does not seem to be always tenable. In fact, in countries that face severe compliance problems, the optimal solution is likely to involve markedly differentiated tax rates on administrative ground alone. This is an important issue and a fertile area for future empirical research.

2.6 Conclusions

This review indicates that the theoretical arguments in favor of uniform indirect taxation for distributional purposes seem rather weak and unrealistic. Almost all empirical-computational studies published so far yield non-uniform optimal tax rates; however, the frameworks of many of these models are significantly different from those underlying the uniform taxation theorems. Many-person computational models consistently yield progressive tax structures. The unresolved controversy concerning commodity tax uniformity calls for more empirical and computational research. Care must be taken when using flexible functional forms in future computational studies, to ensure consistency with theoretical requirements, such as global quasi-concavity of direct utility. The uncritical acceptance of the tax uniformity proposition in some recent policy related studies is a cause for concern. Hopefully, in the future policy advice will rely more on comprehensive and realistic empirical-computational modelling.

---

41 But it should be noted that most empirical-computational models published so far display a fairly low level of complexity. Possible exceptions are: Ebrahim and Heady (1988), Revesz (1997, 2014a, 2014b), Bastani et al. (2014) and Nygard (2014).

42 In other words, a representative consumer model that includes specifications in relation to evasion and administrative/compliance costs, is likely to yield differentiated tax rates depending on these specifications.
Appendix 2.1

Joint optimisation of Mirrleesian income tax with indirect taxation

This appendix will examine how the Mirrleesian income tax formula will change in the presence of proportional commodity taxation. For that purpose, I shall rely heavily on the mathematical analysis presented in Revesz (1989).

In order to prove the change in the left hand side of the optimal marginal income tax formula (discussed in section 2.2.2.5), I shall follow here the steps outlined in appendix A of Revesz (1989). When proportional commodity taxes are present in the model, the Lagrangian defined in (A.2) will be:

\[ L = \int_{w_0}^{w} u[T[x,m(x)], T_m[x,m(x)]] \, dw + \lambda \int_{w_0}^{w} (T[x,m(x)] + \Sigma_j t_j^c p_j q_j(x,m(x))) \, dw - \lambda R \tag{1.1} \]

Here an indirect revenue term \( \Sigma_j t_j^c p_j q_j \) has been added to the direct revenue term (T).

Given fixed commodity tax rates \( t_j^c \), the demand for commodities \( q \) does not depend directly on \( t \) and \( t_m \), but only depends on them indirectly, through changes in \( m \). By applying the \( \delta \) operator variations, a new term is added to (A.12) in Revesz (1989):

\[ \delta t_c^c = \Sigma_j \Sigma_i t_j^c p_j \frac{\partial q_j}{\partial x_i} \delta x_i = \Sigma_j \Sigma_i t_j^c p_j \frac{\partial q_j}{\partial m} \frac{\partial m}{\partial x_i} \delta x_i = \Sigma_j t_j^c p_j \frac{\partial q_j}{\partial m} \delta m = t_m^c \delta m \tag{1.2} \]

Making the \( t_m^c \) term compatible with other derivatives:

\[ t_m^c = \Sigma_i t_i^c p_i \frac{\partial q_i}{\partial m} = \Sigma_i t_i^c p_i \frac{\partial q_i}{\partial w} \frac{\partial w}{\partial m} = \frac{t(c)'}{m'} \delta m \tag{1.3} \]

where \( t(c)' \) is the derivative with respect to \( w \) of total indirect tax revenue.

Using the definition of \( \delta m \) from (A.10) and combining the variations in \( \delta u \) (from A.9), \( \delta t \) (from A.11) and \( \delta t_c \) from (1.2) and (1.3), we obtain the following replacement for (A.12).

\[ \delta L = \int_{w_0}^{w} \left( -\frac{\partial u}{\partial y} \delta t + \lambda \delta t - \frac{\lambda(T + t(c)'w^2)}{(m)^2} \frac{\partial \ell}{\partial w_0} \delta w - \lambda \frac{(T + t(c)')w \partial \ell}{m} \delta t \right) f(w). \tag{1.4} \]

Notice that in (1.4), the two terms \((T' + t(c)')\) replace \( T' \) in (A.12). These two terms consistently replace \( T' \) everywhere in the following steps, leading to the extended version of the optimal marginal income tax formula (eq. (5) in Revesz (2003)), where \( \frac{T_m}{1 - T_m} \) is replaced by \( \frac{T_m + t_m^c}{1 - (T_m + t_m^c)} \) on the left hand side. The fact that there is no explicit change on the right hand side of the marginal income tax formula, suggests that the combined direct plus indirect marginal tax rates in the presence of proportional commodity taxes, will be similar to the optimal solution obtained with income tax alone. Due to possible changes in endogenous variables (incomes, commodities and elasticities), a perfect equivalence in outcomes is unlikely to occur.

In order to make the discussion more explicit, I show here the extended optimal income tax formula, based on the one presented in Revesz (2003) eq. 5.
\[
\frac{T'(m_n) + t_m'}{1 - (T'(m_n) + t_m')} = \frac{d(\ln(m_n))}{dn} \exp\left(\int_{m_0}^m \xi_{(n)}^{R, m_s} ds\right) - \frac{f(n)}{\int_{n_0}^n f(n) \exp\left(\int_{m_0}^{m_s} \xi_{(n)}^{R, m_s} ds\right) dk}
\]

where,
- \(T'(m_n)\) is the marginal income tax rate at wage rate \(n\)
- \(t_m'\) is the marginal indirect tax rate at wage rate \(n\)
- \(n\) is the wage rate (or ability level)
- \(n_0\) is the lowest wage rate for those with positive labour supply
- \(f(n)\) is density of population at wage rate \(n\)
- \(m_n\) and \(m_s\) – gross income at the respective wage rates (note, \(m = n \ell\) where \(\ell\) is labour)
- \(u_c\) marginal utility of income
- \(p\) the Lagrangian, representing the marginal utility of public expenditure
- \(\xi_{(n)}^{R}\) is the compensated elasticity of labour supply
- \(\xi_{(n)}^{R}\) is the derivative of gross income with respect to lump-sum income

This formula applies to a bounded population where the top wage rate (\(\bar{n}\)) is finite. The end point conditions require that \(T'(m_0) + t_m'\) and \(T'(m_\infty) + t_m'\) should be both zero (see Seade (1977)). A similar formula is presented by Saez (2001), in respect to an unbounded population with an asymptotic upper tail.

Now, the question arises, despite no change in appearance, in which situations will the right hand side in (1.5) change in the presence of commodity taxes in the model? An obvious case is when utility is not weakly separable between commodities and leisure. As indicated in section 3.5.6, introducing leisure complements and substitutes into the model leads to a large increase in tax revenue and social welfare. Hence in this case, the combined direct and indirect optimal marginal tax rates will be considerably larger than with income tax alone.

Another interesting case is what happens when we replace part of a pure income tax schedule, by uniform commodity taxes at the rate of \(T_m^0\). In this case we get an indirect tax revenue amounting to \(T_m^0 m\) and a reduction in income tax revenue by the same amount. Theoretically, nothing has changed in respect to utility and demand, and the combined direct plus indirect tax rates remain the same. However, near the end points the marginal income tax rates are reduced to negative values, which appears less politically acceptable than the zero end point conditions that existed before.

As indicated in section 2.2.2.5, eq. (1.5) from above could be used to investigate the effect of some complications in a Mirrleesian joint optimal taxation model through computational studies. Various complexities that affect commodity taxation could be examined, such as tax evasion, externalities, administrative and compliance costs, merit goods as well as non-separable utility between commodities and leisure. But with formula (1.5), complexities related to income taxation could not be investigated, including income tax evasion, administrative and
compliance costs of income tax and concessions in income tax based on certain taxpayer characteristics. It is possible that income tax evasion, as well as administrative and compliance costs, could be incorporated into the optimal variational income tax solution through appropriate enhancements. That seems less likely in respect to different characteristics. I shall not enter into this subject in more detail, because as we shall see in appendix 2.2, there is a simple and robust numerical optimisation method that can quite easily handle all these problems.

Appendix 2.2

Joint optimisation of piecewise linear income tax with indirect taxation

In appendix 2.1 we examined how to carry joint optimisation of income and commodity taxation with a smooth differentiable Mirrleesian income tax function. In this appendix, we shall examine how to carry out joint optimisation with piecewise-linear income tax schedules. As noted in section 2.2.2.2, actual income tax schedules are piecewise linear, hence from a practical point of view, carrying out joint numerical optimisation with these functions may be more policy relevant than optimisations involving a Mirrleesian income tax. Optimisation of a two-bracket piecewise-linear income tax function has been analysed by Apps et al. (2014), together with numerical examples. In this appendix I shall follow a somewhat different approach, using the “gradient” based iterative calculations and the “virtual budget” constraint outlined in chapter 3. In this case, the number of income tax brackets is not restricted to two.

We shall start the discussion with the piecewise-linear income tax defined in the Stiglitz (1982) self-selection model. Later, we turn to the more practical case of continuous piecewise-linear income tax functions. In the Stiglitz (1982) model there are two or more separate income tax brackets, with each bracket having a different marginal tax rate and a different lump-sum tax or subsidy. Denoting the income tax brackets as \((m_0 - m_1) \ldots (m_{N-1} - m_N)\), and assuming that the lump-sum tax/subsidy (denoted \(b_n\)) applies at the start of the bracket, we can apply the gradient optimisation method, described in section 3.3, to the present problem.\(^{43}\)

To recall briefly, the gradient method presented in section 3.3 is applied in the context of a model where no income tax is present, but only commodity taxes and a lump-sum demogrant, representing a particular form of linear income tax. Each iterative approximation uses a given change in a commodity tax rate, which is offset in a revenue neutral manner by an opposite change in the demogrant. In symbolic terms, a tax change of \(\Delta t_j\) is offset by a change in the demogrant amounting to: \(\Delta b = -\Delta t_j p_j Q_j / H\)

\(^{43}\) In analytical models using the Stiglitz self-selection approach, the interest is usually on the effect of a given change on the inclination of high ability workers to mimic the income level of lower ability workers, by reducing their labour supply. We shall not enter into this subject here, and concentrate instead on formal numerical optimisation of the model.
in each iteration, where $Q_j$ is the total quantity purchased from good $j$ and $H$ is the total population.

A similar revenue-neutral method could be applied to find the optimal income tax parameters in each iteration in the self-selection model. A change in the marginal income tax rate of $\Delta T^j_m$ should be offset by changes in all lump-sum taxes/subsidies amounting to

$$\Delta b_j = -\frac{\Delta T^j_m \int_{m_j}^{m_{j+1}} m f dm}{f}$$

for all $b_j$, where $f$ is population density. Note, the term inside the integral is approximately: $H_j(m_j + m_{j+1})/2$, where $H_j$ is the population inside bracket $j$.

Given that in this model there are several lump-sum taxes or subsidies, a commodity tax change should be offset by an opposite changes in all these lump-sum items combined. In this case the revenue-neutral variation is defined for all $b_j$ as:

$$\Delta b_j = -\frac{\Delta f_j Q_j}{H} \quad (2.3)$$

Bearing in mind that in this model, like in the one discussed in section 3.3, $b_0$ is a not independent variable but a dependent variable defined by the budget constraint, we only have to optimize $b_j$s other than $b_0$. In this case, the revenue-neutral equation for a change in $b_i$ will be:

$$\Delta b_0 = -\frac{b_0 H_j}{H_0} \quad (2.4)$$

Given a non-linear income tax, we have to apply the “virtual budget” framework (chapter 3 eqs. 12 and 14) to linearise the net earning parameters at each income level. Having defined the revenue-neutral variations and the net earning parameters, we have most of the information needed to calculate the gradients for all tax variables, using eq. 6 in chapter 3. However, corner point conditions may arise at the boundary points between brackets, where there may be a “bunching” of some taxpayers, who prefer to be located at the boundary, rather than move to an interior point inside the higher or lower tax bracket. These taxpayer groups are not subject to the usual marginal conditions. They could be identified in numerical work, by the following method: First, we establish the point where the taxpayer would be located according to the first order conditions and the linearised net earning parameters (chapter 3, eqs. 12 and 14). After evaluating the utility at that point, we can compare it with the utility he/she would receive at the two adjacent bracket boundaries. Needless to say, the taxpayer would choose the location with the highest utility.

At this stage, a major drawback of the self-selection model should be noted. Since the $b_j$’s are independent variables, the outcome is likely to be a non-continuous income tax function. Given these non-continuities, the convexity of the budget constraint may be violated, and the iterations may not converge to a globally optimal solution. And of course, the shape of the income tax function will not resemble any shape encountered in the real world.

In order to bring the piecewise linear model closer to reality, we have to use continuous income tax functions, without the $b_j$s except for $b_0$ (the demogrant). This form is similar to

---

44 Numerical optimisation of the self-selection model could be used to test the hypothesis raised in Revesz (1989), that a step function type income tax schedule could lead to a higher social welfare than the continuous and differentiable variational solution.
income tax schedules encountered in practice. In this case, we only need to optimize marginal income tax rates apart from commodity taxes. With this type of income tax function, the offsetting change in the demogrant, required following a change in a commodity tax will remain as in (2.1).

When the marginal income tax rate is altered in bracket \( j \) by \( \Delta T_m^j \), even if there is no change in other marginal tax rates, total tax (\( T \)) will have to be adjusted in bracket \( j \) and all higher brackets in order to maintain the continuity of the tax function. Perhaps the best way to solve this problem is to recalculate total tax (\( T^x \)) at each income levels \( m_x \) above \( j \).

For income levels between \( m(j) \) and \( m(j+1) \): \[ \Delta T^x = \Delta T_m^j (m^x - m^j) \] (2.5)

For income levels above \( m(j+1) \): \[ \Delta T^x = \Delta T_m^{j+1} (m^{j+1} - m^j) \] (2.6)

where \( m_x \) is located inside bracket \( k+1 \). Based on the revised \( T^x \) values, we have to recalculate total tax revenue and the demogrant ‘\( b' \). The adjusted \( T^x \) should be also incorporated into “virtual” lump-sum incomes (defined in eq. (14) in chapter 3). Without these adjustments the calculations will be distorted and may not converge to the optimum.

Having obtained new values for \( T_m^j, T^x \) and ‘\( b' \), we can proceed with the gradient calculations (using eq. 6 in chapter 3). Again, bunching at the bracket boundaries may occur, when the marginal income tax rate in one bracket is followed by a higher rate in the next bracket.

As noted earlier, when using the variational income tax solution (eq. 5 in appendix 2.1), it is only possible to handle complexities in relation to commodity taxation but not income tax. No such restriction apply with the continuous piecewise linear income tax model. Provided that demand, utility and labour supply are defined for all relevant combinations of prices and net earning parameters, and the quasi-concavity of direct utility and the convexity of the budget constraint are satisfied, almost any conceivable complication in relation to any tax or taxes, could be properly handled through gradient approximations.45 Apart from the factors analysed in section 3.5 (externalities, tax evasion, non-separable utility, merit goods and administrative and compliance costs), these include a number of other complexities mentioned in the literature, such as:

- Redistributive models with multiple characteristics (Mirrlees, 1976; Akerlof, 1978).
- Different lump-sum grants according to household characteristics (Deaton and Stern, 1986).
- Income uncertainties and their effect on preferences (Cremer and Gahvari, 1995).
- Separate preferences in respect to work participation and work intensity (Saez, 2002b).
- Non-linear commodity tax functions (Atkinson and Stiglitz, 1976; Mirrlees, 1976)

---

45 Of course, there are other possible numerical optimisation methods with piecewise-linear income tax models, apart from the revenue-neutral approach discussed here. For example, we could use the numerical value of the Lagrangian defined in eq. 8b in chapter 3, to calculate the gradient as: \( \lambda \frac{\partial R}{\partial q_i} - u_i \), where \( R \) is total tax revenue and \( u_i \) is the average marginal utility of the consumers of good \( i \). Another possible candidate is the iterative marginal reform method employed in the computational study of Murty and Ray (1989). Dynamic programming could be also used, for optimising piecewise-linear income tax functions in conjunction with commodity taxes.
• Last but not least, multi-generational tax optimisation.

All told, the continuous piecewise linear income tax model offers a robust and highly versatile tool for numerical investigations of joint income-commodity tax optimisation, which can yield more realistic results than extensions of the Mirrleesian model (eq. 5 in appendix 2.1). However, this algorithmic approach is only suitable for numerical analysis, and by itself it cannot provide any new analytical insight. Yet, robust mathematical insights are rare in optimal taxation theory. I think there is little chance for significant progress in this theory, without much more computational work with joint optimisation models. This is a subject that will be discussed in chapter 3.
3. A computational model

3.1 Introduction to chapter 3

This chapter will examine the pattern of optimal commodity tax rates using a computational model. Much of the discussion is concerned with the debate on progressive versus uniform indirect taxation.\textsuperscript{46} The historical background for this debate has been outlined in chapter 2, and will not be repeated here. The chapter is structured as follows:

Section 3.2 examines some recent policy oriented studies on taxation and their relevance to the discussion on differentiated commodity taxation. It also outlines the “growth friendly taxation” approach.

Section 3.3 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 3.2 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. Further details about the structure of the model are outlined in appendix 3.3.

Section 3.4 examines the role of indirect taxation for redistributive purposes. 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. Sub-section 3.4.2 examines the structure of optimal commodity tax rates, in combination with some exogenously given non-linear and non-optimal income tax schedules. Sub-section 3.4.3 deals with a two composite goods extension of the model, which in some respects resembles the tax structure of a number of current VAT systems. Sub-section 3.4.4 reviews some analytical arguments in favour of progressive indirect taxation, mainly on grounds of better work incentives, and the likely sub-optimality of direct support payments. Sub-section 3.4.5 takes a critical look on a proposition in favour of uniform taxation presented by Laroque (2005) and Kaplow (2006). Sub-section 3.4.6 examines how government subsidisation of education and health can fit into the picture.

Section 3.5 deals with a number of factors that can justify differentiated indirect taxation and are not directly related to distributional objectives. These include:

- tax compliance costs
- public administration costs

\textsuperscript{46} We are not dealing in this chapter with non-linear taxes on commodities, which appear in some parts of the theoretical literature.
• indirect tax evasion
• externalities
• paternalistic concerns
• non-separable utility between commodities and leisure

The modified inverse elasticity rule is extended to accommodate these factors, and yields some fairly simple approximate formulas for their impact on optimal tax rates. In some cases these formulas provide reasonably good approximations to the outcomes from iterative calculations. In other cases systematic differences emerge that will be noted, and where possible analysed. The modified inverse elasticity formulas highlight the fact that the impact on optimal tax rates of factors that are seemingly unrelated to distributional objectives, are actually influenced by distributional considerations. Generally, these factors tend to reinforce the progressivity of optimal indirect taxation.

Section 3.6 presents summary and qualifications. Sub-section 3.6.1 summarises the principal findings. Sub-section 3.6.2 examines possible weaknesses in the present model. Sub-section 3.6.3 suggests possible welfare improving and evasion reducing reforms. It recommends to replace part of income tax by higher taxes on non-necessities produced and/or marketed by large organisations, and also to raise property taxes on luxury housing. No consideration is given to the political feasibility of such reforms.

The source code for this computational model is available on the Internet. Appendix 3.1 explains where to find it and how to use it. The interested reader can use this program for numerical experimentations by putting in parameters of his/her choice.

### 3.2 Some recent policy related studies

This section reviews a number of recent policy related studies that were not covered in sufficient depth in chapter 2. These studies examined possible reforms to improve the operation of tax systems. Here we shall review briefly five such studies: Mirrlees et al. (2011), Arnold et al. (2011), OECD (2012) European Commission (2013) and IMF (2014). The emphasis will be on what these studies say in relation to differentiated commodity taxation. With the exception of Mirrlees et al. (2011), this is not a central issue in these reports, nonetheless, they have something to say on this subject.

We start the discussion with the Mirrlees et al. (2011) review of the UK tax system. Among other things, this report recommended to abolish all VAT exemptions and reduced rates that were introduced because of distributional considerations, and replace them by the standard rate. The rationale is that the revenue raised by increasing tax rates on necessities could be used more effectively to provide support to the needy, through reduced income taxes or by increasing welfare benefits.

It is not easy to pass a judgment on this recommendation. To start with, let us note that reducing income tax rates is not a strong option to assist low income households in the UK, because most of them would be on zero or low marginal income tax rates. For that reason, we shall focus here mainly on compensation through welfare payments, whereas Mirrlees et al. (2011) put more emphasis on compensation to the needy through lower income tax. Welfare
payments include pensions, child benefits, unemployment benefits, disability benefits and in-work credits for low wage earners. By using these selective support payments, it is indeed conceivable, although far from certain, that more effective support could be provided to the needy than through zero or reduced VAT rates. Much depends on the administrative effectiveness of expanding these income support schemes. With the possible exception of old age pensions and child benefits, all these support schemes are fairly demanding in terms of monitoring and administration and are vulnerable to false reporting and other misuse. In addition, empirical studies and theoretical considerations (see Revesz (1989)) suggest that unemployment and disability benefits can cause significant work disincentives. Hence, the substitution of lower VAT rates by selective welfare payments is not without its problems. Moreover, given price increases that will affect mainly higher income groups, the overall effect on aggregate welfare is not clear. Nonetheless, there is no denying that such a reform has the potential to be welfare improving. But that does not contradict the arguments presented in this paper in favour of progressive indirect taxation. VAT in the UK (as well as other countries) is far from being a well-developed progressive indirect taxation system (see section 3.4.3). Even in VAT using countries where some progressivity has been retained (such as the UK), items subject to zero or reduced VAT rates cover only a minor share of the value of goods and services sold on the market (European Commission (2003)). To eliminate the minor distributional component of such a system and replace it by targeted income support to the needy, may lead to welfare improvement under specific circumstances. Yet, this takes a rather narrow view on possible reforms. Had the Mirrlees Review posed the question whether it would not be advisable to replace high marginal income tax rates by higher taxes on luxury goods and housing, then the Review would have covered more comprehensively possible reforms in the indirect taxation area. But no such question was raised.

The four other policy studies are all concerned with possible changes in the tax mix, in order to promote “growth-friendly taxation”. All of them recommend replacing at least part of capital and labour taxes by higher taxes on consumption and real estate. European Commission (2013) and IMF (2014) also mention inheritance and gift duties as possible sources for increased revenue, although the importance of these taxes has diminished markedly in recent decades. These studies also recommend to rely more on environmental taxes. The rationale for these recommendations is based on empirical observations that such changes in the tax mix are beneficial for economic development. In particular, Arnold et al (2011) presents econometric evidence, using a longitudinal database of OECD countries, to show that when tax revenue is held constant, consumption and recurrent property taxes have a positive effect on long-term economic growth, whereas the effect of corporate and personal income taxes is negative. In a world of mobile capital, corporate taxes seem to have the worst effects on growth. Generally, the tax treatment of internationally mobile resources (particularly capital), as well as tax havens, has become an increasingly serious concern in recent decades. In order not to get too far away

---

47 Notice that in this substitution exercise we replace a second-best instrument (VAT concessions) by another instrument (selective welfare payments), that at least in theory could approach the effectiveness of a first-best tax/transfer instrument based on ability. Selective welfare payments are discussed in more detail in section 3.4.4.

48 More will be said on present VAT systems from a theoretical perspective in section 3.4.3.
from the main theme, we shall not enter into the various reasons for the empirically observed effects on growth, with one exception.

All the studies agree that as much as possible, labour taxes should be replaced by consumption taxes. The rationale for these tax mix changes is partly based on the disincentive effects of higher marginal income tax rates. No similar disincentives are mentioned in regard to consumption taxes. By implication, these studies perceive income taxation as more detrimental to work incentives than indirect taxation. With the exception of my earlier publications (Revesz (1986, 1997)), I did not see a similar argument appearing in the optimal taxation literature. The less detrimental effect of progressive indirect taxation on labour supply is the main theoretical reason for its advantage over direct taxation, as will be discussed in section 3.4.4. Only Arnold et al (2011) provide some explanation why consumption taxes have a less negative effect on labour supply than income taxes. They reason that since VAT is nearly a flat rate tax, whereas marginal income tax rates are rising with income, income taxation causes more work disincentives. While this may be a valid partial explanation, as we shall see in section 3.4.4 there are also other, perhaps stronger, explanations for the different incentive effects.

With the exception of OECD (2012), all the “growth friendly” taxation studies recommend the replacement of zero or reduced VAT rates by the standard rate (so called “base broadening”). The reasoning is similar to that in Mirrlees et al. (2011), and we shall not discuss it here any further. It is interesting to note that both Mirrlees et al. (2011) and IMF (2014) accept the proposition that there is scope for using progressive indirect taxation in developing countries, where administrative capacity is weak and direct transfer payments are insignificant or non-existent. This idea was raised in the optimal taxation literature in studies dealing with zero or sub-optimal demogrants (see section 2.3.3.1 and 2.3.3.2).

3.3 The mathematical framework of the segmented LES model

The basic structure of the computational study discussed in this chapter, 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 departures 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:

\[ U = \sum_h a_h u_h (q_h) = \sum_h a_h u_h (\mathbf{p}, W_h, y_h) \]  

(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

\[ \sum_h \sum_i t_i \hat{p}_i q_{ih} - Hb - R_0 = 0 \]  

(2a)

and the production possibilities constraint

\[ \sum_h W_h \ell_h - \sum_h \sum_i \hat{p}_i q_{ih} - R_0 = 0 \]  

(2b)

where \( q_i \) are commodities, \( \hat{p}_i \) 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 (the demogrant). Producer prices are fixed. Setting \( \tilde{p}_i \) 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. 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 \( \ell \) is labour supply. Gross income in the post-tax situation after receiving the demogrant (\( b \)) is: 

\[ \hat{m} = W\ell + b. \]

For the time being, lump-sum income ‘\( y_h \)’ equals ‘\( b \)’. Various additions to ‘\( y_h \)’ will be introduced later. Preferences are assumed to be weakly separable, which means that utility is: 

\[ u = u(v(q), L), \]

where \( q \) 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:

\[ u = \sum_i \beta_i \log(q_i - \alpha_i) \]  

The corresponding demand functions for commodities are:

\[ q_i = \alpha_i + \frac{\beta_i}{p_i} (Zw + y - \sum_j p_j \alpha_j) \]  

Labour supply is given as:

\[ \ell = Z - q_L = Z(1 - \beta_L) - \alpha_L - \frac{\beta_L}{w} (y - \sum_j p_j \alpha_j) \]  

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 \( \alpha_i \) can be positive or negative, but with negative \( \alpha_i \) 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.\(^{50}\)

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 \( \beta_i \) parameters of the two groups must be different. Obviously, we are dealing here with non-identical preferences.\(^{51}\)

Appendix 3.3 explains how the utility parameters of the two groups are set. It also explains the definition of the inequality aversion rate used in this model, and its incorporation into the social welfare function weights. The inequality aversion rate ranges from 0 to 1. One represents a high level

\(^{49}\) 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 (\( \ell \)).

\(^{50}\) 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).

\(^{51}\) Whether these should be referred to as heterogeneous preferences or otherwise is discussed in section C.1 in appendix 3.3.
of inequality aversion, corresponding to logarithmic utility, while zero represents the absence of distributional concerns. A few other technical details are also relegated to appendix 3.3.

In regard to the choice of LES as the utility function in the model, let me just note that 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.52

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 (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 that satisfy revenue-neutrality, using constraint (2a). Successive approximations to optimal tax rates are carried out using the gradient equation:

\[ t_i(k) = t_i(k-1) + C \frac{\Delta U_i(k)}{\Delta t_i} \]  

where \( U \) is total utility and \( C \) is a scaling factor determined at the first iteration, \( \Delta U_i(k) = U(t_i(k-1) + 0.03) - U(t_i(k-1)) \) and \( \Delta t_i = 0.03 \). 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 60 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. Mathematical reasoning indicates that given quasi-concave preferences and a convex budget constraint, the optimised, and reports the differences in these totals. The difference in total output is:

\[
\Delta M = \sum_h W_h \Delta \ell_h
\]

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

\[
\Delta U = \sum_h \Delta u_h / \lambda.
\]

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

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

where \( R_0 \) is expenditure on public goods. Changes in \( R_0 \) can be used to estimate the marginal utility of public expenditure.

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 tool to

52 More detailed discussion about utility-demand functions is presented in section 2.3.
53 The convergence proved to be fairly rapid in models without income tax. When income tax is present, the convergence is generally slower and more erratic, which may be due to some possible problem in the program.
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 and its extensions are described in appendix 3.2.

The relevant formula for the model without income tax is (A.11 in appendix 3.2):

\[
t_i \approx \frac{1 + t_i - \bar{u}_i (1 + t_i + \bar{e}_i (t_i - t_{iA}) + \bar{e}_{i, t} \bar{T}_i')}{\bar{e}_i (u_0)}
\]

(9)

The term \(\bar{u}_i = \frac{\partial u_i}{\partial b}\) is called the marginal utility ratio of product i. \(\frac{\partial u_i}{\partial y}\) 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.13 in appendix 3.2). \(\frac{\partial u}{\partial b}\) 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 (see A.2 in appendix 3.2). \(t_{iA}\) is the average commodity tax rate on products other than i. \(\bar{e}_i\) is the consumption weighted average price elasticity of good i, and \(\bar{e}_i (u_0)\) is the corresponding average compensated elasticity of demand. \(\bar{T}_{i, t}\) is the labour supply elasticity term (see A.5a in appendix 3.2) and \(\bar{T}_i'\) is the average marginal indirect tax rate of the consumers of product i (defined in A.5b).

While the modified inverse elasticity rule is not particularly accurate (as illustrated in Table A.1 in appendix 3.2), 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 in (9), 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.: \(\frac{\partial q_j}{\partial p_k} = 0\) for \(j \neq k\). No such restrictive assumption was used in developing the modified inverse elasticity rule, and the two approximations are not the same (see A.14 in appendix 3.2).

When income tax is included, the modified inverse elasticity rule is changed. It will be (A.15 in appendix 3.2):

\[
t_i \approx \frac{1 + t_i - \bar{u}_m (1 + t_i + \bar{e}_i (t_i - t_{iA}) + \bar{e}_{i, t} (\bar{T}_{mi} + \bar{T}_i') - \bar{T}_i (1 - \gamma)}{\bar{e}_i (u_0)}
\]

(11)

where \(\bar{T}_{mi}\) is the average marginal income tax rate of the consumers of product i, \(\bar{T}_i\) is the corresponding average income tax rate, and \(\gamma\) is a general adjustment factor that ensures the equalisation of actual and predicted total tax revenue. \(\gamma\) is usually below plus or minus 0.15, indicating that the optimal combined tax rate \(t_i + \bar{T}_i\) in the presence of income tax, is not much different from the optimal commodity tax rate without it (see appendix 2.1).
3.4 The redistributive model

3.4.1 The model without income tax

In this section we shall examine the basic redistributive model. Various “real life” complexities will be introduced later. The parameters used in all the scenarios reported in this chapter are arbitrary rather than empirical. This approach is perhaps acceptable, given that the book deals with the indirect tax uniformity debate and other broad issues, rather than with the determination of actual tax rates. Moreover, to a large extent the conclusions in this paper are based on analytical arguments, and are unrelated to the parameters used in the numerical calculations. The numerical examples are used to illustrate and substantiate the analytical arguments.

Table 1 shows numerical results from the model where income tax is absent. In the light of the discussion on zero homogenous utility and demand (see section 2.3.1 and C.3 in appendix 3.2), this 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 scenario, utilities are defined by LES without any transformation. Utility transformation, defined by eq. (C.2) in appendix 3.3, occurs in the second 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 (10), 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, \( \bar{u} \)), will usually be larger than the average marginal utility of goods, \( \partial u / \partial y \), because in the later the low marginal utilities 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 drop in demand at the border between the two consumer groups, which results in lower consumption at high income levels (inferior good), leading to the marginal utility ratio of the product to fall 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 for dispersion in the results, besides inequality aversion. In the 0.3 inequality aversion rate scenario, 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 (8a) and the change in output in (7). Notice that the gains in welfare and output over the uniform solution are much larger in the high inequality aversion scenarios. It appears that these gains are positively correlated with the dispersion of optimal tax rates. Gains in welfare and/or output of differentiated 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 the non-uniform than under the uniform solution.

Table 1

<table>
<thead>
<tr>
<th>Inequality aversion = 1</th>
<th>Inequality aversion = 0.3</th>
</tr>
</thead>
<tbody>
<tr>
<td>marginal utility ratios</td>
<td>marginal utility ratios</td>
</tr>
<tr>
<td>compensated elasticity of demand</td>
<td>compensated elasticity of demand</td>
</tr>
<tr>
<td>tax rate</td>
<td>tax rate</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>1</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>-0.45</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
</tr>
<tr>
<td>2</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>-0.79</td>
</tr>
<tr>
<td></td>
<td>0.85</td>
</tr>
<tr>
<td>3</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>-0.78</td>
</tr>
<tr>
<td></td>
<td>0.22</td>
</tr>
<tr>
<td>4</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td>-0.89</td>
</tr>
<tr>
<td></td>
<td>0.99</td>
</tr>
<tr>
<td>5</td>
<td>1.13</td>
</tr>
<tr>
<td></td>
<td>-1.27</td>
</tr>
<tr>
<td></td>
<td>0.48</td>
</tr>
<tr>
<td>6</td>
<td>1.39</td>
</tr>
<tr>
<td></td>
<td>-0.83</td>
</tr>
<tr>
<td></td>
<td>0.98</td>
</tr>
<tr>
<td>7</td>
<td>1.09</td>
</tr>
<tr>
<td></td>
<td>-0.95</td>
</tr>
<tr>
<td></td>
<td>0.45</td>
</tr>
<tr>
<td>8</td>
<td>1.23</td>
</tr>
<tr>
<td></td>
<td>-0.87</td>
</tr>
<tr>
<td></td>
<td>0.68</td>
</tr>
<tr>
<td>9</td>
<td>1.08</td>
</tr>
<tr>
<td></td>
<td>-0.94</td>
</tr>
<tr>
<td></td>
<td>0.43</td>
</tr>
<tr>
<td>10</td>
<td>2.49</td>
</tr>
<tr>
<td></td>
<td>-1.31</td>
</tr>
<tr>
<td></td>
<td>1.49</td>
</tr>
<tr>
<td>11</td>
<td>2.28</td>
</tr>
<tr>
<td></td>
<td>-0.87</td>
</tr>
<tr>
<td></td>
<td>2.68</td>
</tr>
<tr>
<td>12</td>
<td>2.44</td>
</tr>
<tr>
<td></td>
<td>-1.22</td>
</tr>
<tr>
<td></td>
<td>1.62</td>
</tr>
<tr>
<td>13</td>
<td>2.46</td>
</tr>
<tr>
<td></td>
<td>-1.27</td>
</tr>
<tr>
<td></td>
<td>1.55</td>
</tr>
<tr>
<td>14</td>
<td>2.57</td>
</tr>
<tr>
<td></td>
<td>-1.48</td>
</tr>
<tr>
<td></td>
<td>1.32</td>
</tr>
<tr>
<td>15</td>
<td>2.34</td>
</tr>
<tr>
<td></td>
<td>-1.00</td>
</tr>
<tr>
<td></td>
<td>2.12</td>
</tr>
<tr>
<td>16</td>
<td>2.32</td>
</tr>
<tr>
<td></td>
<td>-0.94</td>
</tr>
<tr>
<td></td>
<td>2.34</td>
</tr>
<tr>
<td>17</td>
<td>2.58</td>
</tr>
<tr>
<td></td>
<td>-1.5</td>
</tr>
<tr>
<td></td>
<td>1.31</td>
</tr>
<tr>
<td>18</td>
<td>2.42</td>
</tr>
<tr>
<td></td>
<td>-1.17</td>
</tr>
<tr>
<td></td>
<td>1.70</td>
</tr>
</tbody>
</table>

Average tax rate on the 9 necessities 0.70 0.27
Average tax rate on the 9 luxuries 1.83 0.41
% average tax from expenditure – non-uniform 48.4 23.8
% average tax from expenditure –uniform 50.8 25.7
% demogrant to average labour income 69 17
% change in welfare terms compared to uniform solution 3.5 0.2
% change in total output compared to uniform solution 2.1 1.9

* Fixed expenditure on public goods = 10% of total income under zero tax
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 in section C.3 in appendix 3.3, in a basic redistributive model (no administration and evasion) zero income tax combined with a demogrant, 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.\(^54\) Among earlier computational studies, I think only Ebrahimi and Heady (1988) 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 (1988) indicate that under weakly separable utility, optimal commodity tax rates will be differentiated, provided Engel curves are not parallel across households. Non-parallel Engel curves lead to some departure from strict linearity. Even that slight non-linearity of Engel curves has led to a perceptible differentiation in optimal tax rates.\(^55\)

### 3.4.2 The model with a non-linear income tax

So far we discussed the structure of optimal commodity tax rates under a linear income tax. At this stage the question arises, as 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 when utility is weakly separable, then the optimal non-linear income tax (Mirrlees (1971)) should be combined with zero or 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 income 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.

The mathematical framework is changed after the introduction of income tax. Following the “virtual budget” framework outlined in Revesz (1986, 1997), Roberts (2000) and Saez (2001), the earning parameters used in (4) and (5) have to be redefined. The post-income-tax wage rate is:

\[
    w = W(1 - T_m) \tag{12}
\]

where \(T_m\) is the marginal income tax rate and \(W\) is the pre-income-tax wage rate (or ability level). Post-income-tax ‘virtual’ lump-sum income (\(y\)) is defined so that \(w + y\) add up to disposable income (\(m\)):

\[
    m = t W - T + b = t w + y \tag{13}
\]

\(^54\) The approximate formula for optimal tax rates developed by Sandmo (1975) also suggests that in a model without income tax, optimal tax rates tend to be progressive. This formula is presented in appendix 3.2 eq. (A.14).

\(^55\) 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.
From (12) and (13) it follows that post-income-tax “virtual” lump-sum income is:

\[ y = T_m W_t - T + b \]  

(14)

In (14) pre-tax lump-sum income \((Y_0)\) is assumed to be zero. Note, in the presence of income tax, the linearised earning parameters \(w\) and \(y\) are the dual counterparts to income and leisure in the indirect utility space. The “virtual” lump-sum income component \((y)\) depends on the curvature of the income tax function \((T)\). It will equal ‘b’ with a linear income tax. It will be above ‘b’ if marginal income tax rates are increasing, and vice versa if they are decreasing.

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). In the following discussion, when we investigate scenarios where income tax is present, it will be always the progressive schedule described above. The regressive schedule is mainly of theoretical interest.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Progressive schedule</th>
<th>Regressive schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inequality aversion=1</td>
<td>Inequality aversion=0.3</td>
</tr>
<tr>
<td>Average tax rate on necessities</td>
<td>0.22</td>
<td>-0.16</td>
</tr>
<tr>
<td>Average tax rate on luxuries</td>
<td>0.91</td>
<td>-0.13</td>
</tr>
<tr>
<td>% demogrant over average income income</td>
<td>42</td>
<td>-2</td>
</tr>
<tr>
<td>% change over uniform in welfare terms</td>
<td>2.8</td>
<td>0.4</td>
</tr>
<tr>
<td>% change over uniform in total output</td>
<td>3.0</td>
<td>0.6</td>
</tr>
<tr>
<td>Marginal tax rates</td>
<td>Marginal tax rates</td>
<td>Marginal tax rates</td>
</tr>
<tr>
<td>bracket 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>bracket 2</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>bracket 3</td>
<td>0.1</td>
<td>0.2</td>
</tr>
<tr>
<td>bracket 4</td>
<td>0.3</td>
<td>0.3</td>
</tr>
<tr>
<td>bracket 5</td>
<td>0.4</td>
<td>0.4</td>
</tr>
</tbody>
</table>

* Fixed expenditure on public goods = 10% of total income under zero tax

In three out of the four scenarios in Table 2, 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 negative and nearly uniform commodity taxes (see Table D.1 in appendix 3.4). When the inequality aversion rate 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 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), indicate 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 section C.1 in appendix 3.3, in a many-goods model linear Engel curves for all goods represents nearly homothetic preferences, which does not accord with empirical evidence. For econometric evidence against linear Engel curves for all goods refer to Blundell and Ray (1984).

The preliminary results in Table 2 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 non-linear income tax schedules which by themselves almost satisfy distributional objectives, and where the associated optimal commodity tax rates are negative – in that event the solution tends to move towards uniformity. This can only happen if a non-linear income tax schedule is included in the model.

3.4.3 The two composite goods model

Value added tax (VAT) is nowadays the most widely used indirect tax system in developed countries. Therefore, it is of some interest to examine these taxes in the light of the numerical model under discussion. In many countries VAT is nearly a flat rate tax, with little progressivity built into it. In the following discussion we shall focus on VAT systems where some progressivity has been retained, such as in the UK, France, Italy, Spain, Ireland and Luxembourg (see European Commission (2014)). VAT in these countries can be characterised in broad terms as a two tiered tax system. Most goods and services are taxed at the standard VAT rate or close to it. A minority of goods (mainly necessities) are taxed at zero or significantly reduced rates. These include: food, medicines, private medical and dental care, passenger transport, rented accommodation, books and periodicals and in some countries also domestic fuel, social housing, and children clothing.

To describe these systems in modelling terms, we used a two composite goods approach, where the luxury good is taxed at the standard VAT rate and the necessity is taxed at a much reduced rate. This approach is suitable to examine also indirect tax systems other than VAT (such as sales tax), where separate tax rates are defined for very broad categories of goods and services. Surely, the assumption used earlier, that 18 separate tax rates can be established for 18 goods is unrealistic. All legal tax codes are based on broad definitions of product groups. Arguably, in the computer age it might be possible to administer indirect tax systems where separate tax rates are determined for individual products based on quality, using measures such as price-weight or price-volume ratios or property valuations, but that is a vision for the future, far removed from current practice.

In the two composite goods model, the original 18 goods have been retained, but the utility parameters of all nine necessities have been set to the same values and another set of identical values was applied to the nine luxuries (see Table C.2 in appendix 3.3). Hence, effectively there are only two commodities in the model.
In order to understand better “progressive” VAT systems, I searched for a solution where the tax on luxuries is 23% and the tax on necessities is 2%. It was not easy to find such a solution. Usually the differences between optimal tax rates in the two-goods model are much larger than those between these target figures. But eventually I found a not too realistic combination that yielded these figures as optimal outcomes. This combination corresponds to no income tax, an inequality aversion rate of 0.13 and fixed expenditure on public goods amounting to 22% of total pre-tax output. Under these specifications, given 2% and 23% commodity tax rates, the demogrant turned out to be minus 10% of average income. The first column in Table 3 shows the optimal solution under these specifications, the other columns show the solution under higher inequality aversion rates and with progressive income tax.

<table>
<thead>
<tr>
<th>Without income tax</th>
<th>With progressive income tax</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public goods = 22%</td>
<td>Inequality aversion rates</td>
</tr>
<tr>
<td>0.13</td>
<td>0.3</td>
</tr>
<tr>
<td>Average tax rate on the necessity</td>
<td>0.02</td>
</tr>
<tr>
<td>Average tax rate on the luxury</td>
<td>0.23</td>
</tr>
<tr>
<td>% demogrant to average income</td>
<td>-10</td>
</tr>
<tr>
<td>% change in welfare terms compared to uniform solution</td>
<td>0.2</td>
</tr>
<tr>
<td>% change in total output compared to uniform solution</td>
<td>0.8</td>
</tr>
</tbody>
</table>

The figures in Table 3 illustrate a few points. First, in the present model even the more “progressive” VAT systems can be justified only under a very low inequality aversion rate, in fact lower than those used elsewhere in this chapter. But it should be noted that the VAT estimates used here exclude other indirect taxes, such as excises and property taxes. Under higher inequality aversion rates (0.3 and 1), the dispersion of optimal tax rates in the two-goods model is considerably larger than the dispersion of tax rates in the 18 commodities model, reported in Tables 1 and 2. My original expectation was that broad categorisation of taxable items will reduce the dispersion of optimal tax rates. The optimisation results point in the opposite direction. But some caveats should be noted. By construction, segmented LES utility draws a sharp distinction between necessities and luxuries, which enables a clear-cut partitioning in the model between these categories. The broad legal definitions of products for taxation purposes, which ignore qualitative differences between products within the same group, do not allow for similar well defined distinctions between luxuries and necessities. Hence in practice, separate tax rates for broad product groups will not offer as much advantage over uniform taxation as suggested in Table 3.

Given that “progressive” VAT is not much different from a flat rate tax, according to the present model transforming it into uniform taxation will reduce welfare by only 0.2%. But
Because lump-sum income discourages work, the opposite direction. Looking at (12) and (14), higher $T_m$ will reduce $w$ and increase $y$, which lower $y$. But higher marginal income tax rates move the earning parameters exactly in the opposite direction. We start the discussion with the linearised after-income-tax earning parameter defined in (12) and (14). The incentive effects of $w$ and $y$ can be seen from the utility compensated labour supply derivative.

$$\frac{\partial \ell(U_o)}{\partial w} = \frac{\partial \ell}{\partial w} + \frac{\partial \ell}{\partial y} \frac{\partial y(U_o)}{\partial w} = \frac{\partial \ell}{\partial w} - \frac{\partial \ell}{\partial y} \frac{\partial y}{\partial w} \frac{\partial \ell}{\partial y}$$  \hspace{1cm} (15)

Because lump-sum income discourages work, $\partial \ell/\partial y$ is always negative. The compensated derivative for labour must be definite positive, because labour is a negative good. From (15), this implies that utility compensated variations that increase labour supply require higher $w$ and lower $y$. But higher marginal income tax rates move the earning parameters exactly in the opposite direction. Looking at (12) and (14), higher $T_m$ will reduce $w$ and increase $y$, which means that increasing marginal income tax will discourage work. On the other hand, with commodity taxation, regardless whether progressive or regressive, both $w$ and $y$ will change in the same proportion, determined by the change in the perfect consumer price index. Unlike with progressive income taxation, there is no shift that reduces $w$ and increases $y$. That means that replacing progressive commodity taxation by an equal revenue generating progressive income tax will reduce labour supply, because following such a change $w$ will decrease and $y$ will increase.\footnote{The main source for disincentives in the commodity tax model is the demogrant, which being lump-sum income will reduce labour supply. The demogrant also reduces labour supply in income tax models, in addition to other disincentives.} This had been confirmed by numerical results from an earlier version of the present computational model reported in Revesz (1997).\footnote{These calculations are also included in the current version of the program, but are not reported here in order to save on space.} Table 3 in Revesz (1997) shows that a changeover from progressive indirect taxation to an equal revenue generating income tax schedule, will cause a large decrease in total output and welfare. These results, based on standard consumer theory, confirm the implicit assumption in the “growth friendly taxation” reports (see section 3.2), that replacing income taxes by consumption taxes will improve work incentives.

But there is more to it. The tax burden due to income taxation is more visible than that due to commodity taxation. On behavioural grounds, one would expect that the response to a more visible disincentive will be stronger than the response to a less visible one. This behavioural response also supports the view, that progressive direct taxation will have a more detrimental effect on incentives than progressive indirect taxation. The fact that with progressive indirect taxation substantial redistribution can be carried out, with little detrimental effect on labour supply, is the prime reason why the optimal indirect tax solutions reported in this chapter tend to be progressive rather than uniform or regressive.

3.4.4 Analytical arguments in favour of progressive indirect taxation

At this stage the question arises, why do we get progressive commodity taxes in the optimal solutions instead of uniform taxes? As we shall soon discover, to a large extent the answer is connected with labour supply incentives, even when preferences are weakly separable between commodities and leisure. We start the discussion with the linearised after-income-tax earning parameter defined in (12) and (14). The incentive effects of $w$ and $y$ can be seen from the utility compensated labour supply derivative.

$$\partial \ell(U_o)/\partial w = \partial \ell/\partial w + \partial \ell/\partial y \partial y(U_o)/\partial w = \partial \ell/\partial w - \partial \ell/\partial y \partial y/\partial w \partial \ell/\partial y$$  \hspace{1cm} (15)

Because lump-sum income discourages work, $\partial \ell/\partial y$ is always negative. The compensated derivative for labour must be definite positive, because labour is a negative good. From (15), this implies that utility compensated variations that increase labour supply require higher $w$ and lower $y$. But higher marginal income tax rates move the earning parameters exactly in the opposite direction. Looking at (12) and (14), higher $T_m$ will reduce $w$ and increase $y$, which means that increasing marginal income tax will discourage work. On the other hand, with commodity taxation, regardless whether progressive or regressive, both $w$ and $y$ will change in the same proportion, determined by the change in the perfect consumer price index. Unlike with progressive income taxation, there is no shift that reduces $w$ and increases $y$. That means that replacing progressive commodity taxation by an equal revenue generating progressive income tax will reduce labour supply, because following such a change $w$ will decrease and $y$ will increase.\footnote{The main source for disincentives in the commodity tax model is the demogrant, which being lump-sum income will reduce labour supply. The demogrant also reduces labour supply in income tax models, in addition to other disincentives.} This had been confirmed by numerical results from an earlier version of the present computational model reported in Revesz (1997).\footnote{These calculations are also included in the current version of the program, but are not reported here in order to save on space.} Table 3 in Revesz (1997) shows that a changeover from progressive indirect taxation to an equal revenue generating income tax schedule, will cause a large decrease in total output and welfare. These results, based on standard consumer theory, confirm the implicit assumption in the “growth friendly taxation” reports (see section 3.2), that replacing income taxes by consumption taxes will improve work incentives.

But there is more to it. The tax burden due to income taxation is more visible than that due to commodity taxation. On behavioural grounds, one would expect that the response to a more visible disincentive will be stronger than the response to a less visible one. This behavioural response also supports the view, that progressive direct taxation will have a more detrimental effect on incentives than progressive indirect taxation. The fact that with progressive indirect taxation substantial redistribution can be carried out, with little detrimental effect on labour supply, is the prime reason why the optimal indirect tax solutions reported in this chapter tend to be progressive rather than uniform or regressive.
Another point in favour of progressive indirect taxation is connected with the modelling of income support. In almost all 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)). Given that support agencies have access to more information on recipients than tax authorities on taxpayers, it is theoretically possible to get closer to providing support based on abilities than to tax according to abilities. As explained in Mirrlees (1971), taxing or subsidising on the basis of abilities is the first-best optimal solution to the distributional problem, but in practice abilities (i.e. fixed endowments or potential) cannot be properly observed. In any event, provided income support is organised and monitored efficiently, the provision of targeted welfare payments stands closer to the ideal of ability-based transfers than either income or commodity taxation. The effectiveness of differentiated and targeted support payments to the needy is of central importance in actual tax-transfer 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 taxation models, as explained in sections 2.2.2.4 and 2.3.3.2. Since most income support systems have to screen partly on the basis of semi-observable characteristics, they will fall short of the ideal of providing differentiated lump-sum grants based on actual family subsistence needs and ability to work. In this situation, progressive indirect taxation can be used to compensate for some of the inherent shortcomings in the income support system. These benefits from progressive indirect taxation are not reflected in the present numerical model, which assumes an optimal uniform demogrant.

Finally, we come to the subject of sub-optimal demogrants, which may also provide a rationale for progressive indirect taxation, particularly in developing countries, as indicated in section 2.3.3.2. As noted there, the models of Ray (1989) and Asano et al. (2004) assume sub-optimal demogrant, rather than zero demogrant. Numerical simulations with the present model, presented in table D.2 in appendix 3.4, suggest that when an exogenously given demogrant is set below the endogenously determined optimal level, then optimal commodity tax rates will be more progressive the further away the fixed demogrant is located from the optimum. Progressivity is measured by the ratio of the average tax on luxuries over the average tax of necessities. When the fixed demogrant is set to zero, there is a large increase in indirect tax progressivity compared with the endogenously determined optimal solution.

Yet despite many modelling possibilities, there is a fundamental problem with this approach. To what extent a given demogrant is sub-optimal cannot be quantified in an objective manner. It all depends on the eyes of the beholder, or in mathematical terms, on the utility function and inequality aversion rate employed in the model. Moreover, the situation in developing countries can be represented also in a model where the demogrant is determined endogenously, by taking out income tax and incorporating into the model high evasion propensities for many products, high administrative costs and low inequality aversion rate. Experiments with the present model along these lines, yielded relatively high tax rates on less
evasion prone items and marked progressivity in tax rates. Hence, in order to justify progressive indirect taxation in developing countries, there is no need to assume a sub-optimal demogrant. Nonetheless, in many countries where a large percentage of the population are working in the “informal economy”, it is impossible to generate sufficient revenue for redistribution, in line with the preferences of the median voter or the ruling elite (De Freitas (2012)). In order to model such a situation, the sub-optimal demogrant approach has its merits. If one accepts it as a valid argument, then it can be added to a number of other arguments justifying progressive indirect taxation.

3.4.5 Discussion on 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 (L-K) 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 L-K proposition extends the Atkinson-Stiglitz theorem to non-optimal income tax functions as well.

At first sight the L-K 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, be it linear or non-linear. 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 the discussion. Moreover, the L-K 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 L-K proposition.

At this stage, we could leave the discussion on the L-K proposition with this inconclusive statement, however, because the L-K 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 L-K proposition has been presented by Boadway (2010). Among other things, Boadway points out that the L-K 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 L-K 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 L-K proposition.
For the purpose of analysing this issue, we look at a simple corollary to the L-K 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 Seade (1977), 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. As indicated in section 2.2.2.2 such an income tax schedule is not politically acceptable in a world of imperfect competition and information. That implies that the L-K reform may sometimes (but not always) steer toward an income tax solution that is politically unacceptable. Consequently, the warning by Boadway (2010) about the possible infeasibility of carrying out the income tax adjustment required according to the L-K 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.

The “real life” complexities mentioned in section 2.2.2.2 to 2.2.2.4, invalidate the Atkinson-Stiglitz theorem as well as the L-K proposition. Another practical objection relates to the L-K 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 the political acceptability of flexible adjustments to income tax is fairly limited. This could be another impediment to the practical application of the L-K proposition. 58

58 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-tax8 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.
Apart from these practical application issues with a highly stylised and simplistic model, there is also a theoretical problem with the L-K model connected with Pareto improvement. To prove Pareto improvement, the authors assume that the first step in the L-K 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 L-K reform in this simple model.

A crucial assumption in the L-K 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 $U = f(v(c), \ell)$, 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 $\ell$ remains constant so will $U$. Hence, constant $\ell$ and $U$ is a feasible outcome. However there are another two possible outcomes associated with constant $v$ changes – either increase $\ell$ and reduce $U$ or decrease $\ell$ and increase $U$. It is not difficult to see that the option to increase $\ell$ and reduce $U$ will not be a rational choice. That still leaves open either constant $\ell$ and $U$ or decreasing $\ell$ and increasing $U$.

The possible choice of decreasing $\ell$ 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 $\ell$ 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 $\ell$ 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 ($y$) discussed in section 3.4.2. To recall briefly: the net wage rate is defined as: $w = W(1 - T')$ (12) and “virtual” lump-sum income as: $y = T'w - T$ (14) $T$ represents total income tax and $T'$ is the marginal income tax rate. The labour supply function is given as $\ell = f(w, y, p)$. 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 tax, labour supply of at least a section of the population is reduced, then
government revenue might be curtailed following the L-K 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 \( \ell \) 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, \( Y= W\ell \). Where Laroque (2005)
probably made a mistake, is in assuming that having access to the same menu (expressed in
terms of \( W\ell \)) will ensure the same labour choice, despite possible changes in the net earning
parameters \( w \) and \( y \), which are completely 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 \( (\ell) \) remain unchanged. Kaplow (2006, p. 1240) defines
\( T^0(W\ell) \) so that:

\[
v(t, T, W\ell) = v(t^*, T^0, W\ell) \quad \text{for all } W\ell.
\]

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 (12)
and (14), following the replacement of \( T \) by \( T^0 \) labour supply is given as:

\[
\ell(w, y, p) = \ell(W(1 - T^0''), (T^0'W \ell - T^0), p).
\]

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 :

\[
\ell_0'(W) = \ell(T^0', T^0, p)
\]

where \( T^0'' \) is the derivative of \( T^0 \) with respect to \( W\ell \) and \( \ell_0 \) is pre-reform labour supply. The
likely mistake in Lemma 1 (p. 1241) occurs in the second equation, where Kaplow claims that:

\[
U(v(t, T, W\ell), \ell) = U(v(t^*, T^0, W\ell), \ell) \quad \text{for all } W\ell.
\]

Kaplow explains that this equality follows from (16). In (18) the left hand side refers to the
pre-reform situation and the right hand side to the post-reform situation. As explained
earlier, in (18) \( \ell \) on the left hand side would equal to \( \ell \) on the right hand side, only if the
differential equation in (17) is satisfied. But there is no reason to believe that generally the same
\( T^0(W\ell) \) function will satisfy both conditions (16) and (17), which leads us to conclude that \( \ell \)
on the two sides of (18) 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. The main problem with their analysis is that they ignored completely the “virtual
budget” framework, which explains the different disincentive effects of income and commodity
taxation.
3.4.6 Subsidies for education and health

So far, when we referred to commodities, we tacitly assumed that education and health are excluded from this category. But in economic terms education and health are services like all other services, and are unique only in the sense of being almost fully subsidised by the government. Even without government support, there would be considerable demand for these services through the market. The economic rationale for government subsidisation of these services is based on the same principles that justify differentiated commodity taxes/subsidies in general. These are:

- Distributional concerns
- Paternalistic concerns
- Externalities – there are positive externalities associated with education, through the unpaid diffusion of knowledge to others. In health positive externalities arise through the prevention of infectious diseases.
- Non-separable utility between commodities and leisure. Education is a work complement because it raises the quality of labour supply. Health is a work complement because it increases the size of the working population.

In both education and health, distributional concerns are one of the main motives for public funding (see Hindriks and Myles (2006)). In view of these facts, those who argue that there is no need for differentiated indirect taxation for distributional purposes, if they wanted to maintain logical consistency, should also advocate the reduction of government expenditure on education and health, in order to use these funds to reduce income taxes and/or increase welfare payments. I have not seen such a suggestion in the optimal taxation literature, which indicates that perhaps the proponents of uniform commodity taxation are not aware to the full implications of their position.

3.5 Other factors justifying differentiated indirect taxation

In this section we shall examine a number of other factors that can justify differentiated commodity taxation, and are not directly related to distributional concerns. These include: compliance costs, government administration costs, tax evasion, externalities, paternalistic concerns and non-separable utility between commodities and leisure. To assess analytically the impact of these factors, a common methodology is used, based on the modified inverse elasticity rule. We assume that each factor can influence optimal tax rates predicted by the modified inverse elasticity rule, either through changes in the lump-sum incomes of consumers or changes in tax revenue or both. We examine how the formula will change following the introduction of a factor. Finally, we deduct from the revised formula the original formula, to obtain an analytical expression for the net change in predicted tax rates due to the factor concerned. In following this estimation method of net impact, we ignore possible changes in endogenous variables, such as demand elasticities, labour supply elasticities, marginal utilities, other tax rates and the demogrant. Despite the strong assumptions adopted, the numerical results suggest that these analytical approximations of net impact work reasonably well. In each of the following sections we shall compare the analytical approximation with numerical results from iterative calculations.
3.5.1 Compliance costs

We start the discussion with tax 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:

\[
C_i = c_i p_i q_i t_i
\]

Taking the indirect utility definition of the social welfare function in (1), then the dead-weight costs defined in (19) can be incorporated into the social welfare function by subtracting them from personal lump-sum incomes \('y'\). The extended social welfare function will be:

\[
U = \sum_h a_h u_h(P, W, [y_{oh} - \sum_i c_i p_i q_i t_i])
\]

Using an extension to the modified inverse elasticity rule in appendix 3.2, we derive in eq. (A.19) the following approximation for the impact of compliance costs on optimal tax rates.

\[
\Delta t_i(c_i) \approx \frac{[1 + t_i (1 + \bar{e}_i)] c_i}{\bar{e}_i(0)}
\]

Assuming that \(\bar{e}_i\) and \(\bar{e}_i(0)\) are the same and equal to -1 (see Tables 1 and 4), we can obtain from (21a) the cruder but simpler approximation:

\[
\Delta t_i(c_i) \approx - c_i
\]

Table 4

<table>
<thead>
<tr>
<th>Good</th>
<th>necessity</th>
<th>luxury</th>
<th>Good</th>
<th>necessity</th>
<th>luxury</th>
<th>Good</th>
<th>necessity</th>
<th>luxury</th>
<th>Good</th>
<th>necessity</th>
<th>luxury</th>
<th>Good</th>
<th>necessity</th>
<th>luxury</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>inequality aversion</td>
<td>With progressive income tax</td>
<td>Without income tax</td>
<td></td>
<td>Demand elasticities (\bar{e}_i)</td>
<td>(\bar{e}_i(0))</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 5</td>
<td>1</td>
<td>0.03</td>
<td>0.05</td>
<td>-0.02</td>
<td>-1.25</td>
<td>-1.17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 18</td>
<td>1</td>
<td>0.80</td>
<td>0.83</td>
<td>-0.03</td>
<td>-1.17</td>
<td>-1.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 2</td>
<td>0.3</td>
<td>-0.13</td>
<td>-0.13</td>
<td>0.00</td>
<td>-0.81</td>
<td>-0.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 12</td>
<td>0.3</td>
<td>-0.15</td>
<td>-0.15</td>
<td>0.00</td>
<td>-1.11</td>
<td>-1.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Having obtained a theoretical approximation, we can now look in Table 4 at some numerical results. These results were obtained by taking the basic model with or without income tax, and adding a compliance cost parameter of 4% to two goods, one a necessity the other a

---

59 Note the small differences between \(\bar{e}_i\) and \(\bar{e}_i(0)\) in Table 4. With LES utility the difference between the two is \(\bar{\beta}_i\). Given the condition that \(\sum_i \bar{\beta}_i = 1\), then with 18 goods plus leisure that means that on the average \(\bar{\beta}_i\) is below 0.06, and so is the difference between \(\bar{e}_i\) and \(\bar{e}_i(0)\).
luxury. The resulting $t_i$’s are then compared with the original tax rates obtained without compliance costs, shown in Table 1 and Table D.1 in appendix 3.4.

With negative taxes the impact of compliance costs is zero. The reason is that when taxes are negative, the program sets compliance, administration and evasion costs to zero, because it is hard to put an interpretation to these costs in the presence of subsidies. Otherwise, the results appear to be in line with what is expected according to (21b). Overall, the results indicate that compliance costs will reduce optimal tax rates by a similar amount to the share of these costs from commodity taxes.

3.5.2 Government administration costs

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 $s$, public revenue will be affected.

Define the total cost of public tax administration as: $S = \sum_h \sum_l p_l q_{ih} t_i s_i$ (22)

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

$$R = \sum_h \sum_l (p_l q_{ih} t_i - p_l q_{ih} t_i s_i) - R_0$$ (23)

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

The incidence of administrative costs on consumers or the government, has an effect on the optimal tax rate. The estimated impact of government administration according to the modified inverse elasticity rule (A.23 in appendix 3.2) will be:

$$\Delta t_i(s_i) \approx \frac{s_i \bar{u}_{mi} \left[1 + t_i(1 + \bar{e}_i)\right]}{\bar{e}_i(u_0)}$$ (24a)

Again, assuming that $\bar{e}_i$ and $\bar{e}_i(u_0)$ equal to -1 (see Tables 1 and 4), we obtain from (24a) the cruder but simpler approximation:

$$\Delta t_i(s_i) \approx -s_i \bar{u}_{mi}$$ (24b)

The notable difference between the approximations in (21) and (24) is the multiplication of the administrative cost term in (24) 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 impact of government administration costs on optimal tax rates is expected to be larger than the impact of identical compliance costs.

Table 5 presents results for the same products reported in Table 4. These figures suggest that the impact of administration costs on optimal tax rates is indeed somewhat larger than those of compliance costs, in line with what is predicted from (21) and (24). Like with compliance, administration costs will reduce optimal tax rates by a similar amount to the share of these costs from commodity taxes.
Another important variable is the dead-weight costs of evasion – its constant 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 a particular good. Apart from concealment costs, there are sometimes even larger dead-weight losses associated with less efficient production methods employed in smaller less organised firms in the black or grey economy. Taking these definitions, total evasion amounts to

$$E_i = e_i q_i t_i$$  \hspace{1cm} \text{(25)}$$

and the dead-weight cost of evasion ($D_i$) is given as:

$$D_i = e_i d_i q_i t_i$$  \hspace{1cm} \text{(26)}$$

where $q_i$ is total expenditure on good $i$.

Definition (25) is very similar to (19) on compliance and (22) on administration. Consequently, we shall model evasion by synthesising the compliance and administration models. In my earlier papers (Revesz (1997, 2014a)), I presented an evasion model where the loss in government revenue is partly offset by effective (post-evasion) price reductions, and these will benefit all the consumers of the product. However, there are some empirical weaknesses in the assumption that evasion will necessarily be reflected in lower effective prices, and these will benefit all the consumers of the product. Arguably, a large part of evasion is not translated into lower effective prices, but instead brings income gains to a small group of

---

60 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). Cremer and Gahvari (1993), cited in section 2.2.2.3, also follow this approach.
tax evaders. For that reason, we shall develop here a simpler model of evasion that does not involve effective price reductions.

It is assumed that evasion on product i causes $E_i$ loss to government revenue. The associated dead-weight costs are defined in (26). Excluding these dead-weight costs, the loss in government revenue brings benefits to some consumers. In one scenario we assume that the benefits of evasion are spread among the consumers of product i in direct proportion to consumption.

Alternatively, we use predetermined distribution weights (denoted $g_{ih}$) to allocate benefits to consumers. By definition:

$$\sum_h g_{ih} = 1 \quad \text{and} \quad n_{ih} = (1 - d_i)E_i g_{ih}$$

(27)

where $n_{ih}$ represents the gain to taxpayer h from evasion associated with good i.

The change in the modified inverse elasticity rule due to evasion will have two components. One is the loss in government revenue, which follows the same formula as government administration costs in (24b), with $s_i$ replaced by $e_i$. When the gains are spread among the consumers of the product through higher lump-sum incomes, these gains will be of opposite sign to the lump-sum costs of compliance in (19), with $c_i$ being replaced by $-(1 - d_i)e_i$.

Using (21b) and (24b), the approximate impact of evasion on the optimal tax rate is given as:

$$\Delta t_i(e_i) \approx (1 - d_i)e_i - e_i \bar{u}_{mi}$$

(28a)

When the benefits are spread according to the distribution weights $g_{ih}$, then the average marginal utility of recipients will be:

$$\bar{u}_{gi} = \sum_h \frac{\partial u_h}{\partial y} g_{ih} \, .$$

(29)

In this case the impact on optimal tax rate is approximated as:

$$\Delta t_i(e_i) \approx (1 - d_i)e_i \frac{\bar{u}_{gi}}{\bar{u}_i} - e_i \bar{u}_{mi}$$

(28b)

where $\bar{u}_i = \frac{\partial u_i}{\partial y}$ is the average marginal utility of the consumers of product i, defined in (A.13) in appendix 3.2. Here the benefits to consumers are valued according to $\bar{u}_{gi}$ instead of $\bar{u}_i$. If the benefits accrue mainly to high income earners, then unless $q_i$ is a strong luxury, $\bar{u}_{gi} < \bar{u}_i$ and the decrease in the optimal tax rate due to evasion will be larger than under (28a). If the benefits accrue mainly to low wage earners the opposite applies.

Now, let us look at some numerical results. Table 6 is structured the same way as Tables 4 and 5, and the iterative calculations follow the same procedure. It is assumed in these calculations that evasion gains are spread among the consumers of the product, therefore in this case (28a) applies.

---

61 There is no need to assume that all consumers will benefit from evasion. As long as the gains are spread randomly within the group, the average marginal utility of beneficiaries will be the same as that of the consumer group as a whole.
Table 6

OPTIMAL TAX RATES WITH COMMODITY TAX EVASION

<table>
<thead>
<tr>
<th>Product No. and type with evasion</th>
<th>without evasion</th>
<th>impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Without income tax, inequality aversion = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 4 - necessity</td>
<td>0.59</td>
<td>0.99</td>
</tr>
<tr>
<td>Good 14 - luxury</td>
<td>1.16</td>
<td>1.32</td>
</tr>
<tr>
<td>Good 4 - necessity</td>
<td>0.12</td>
<td>0.32</td>
</tr>
<tr>
<td>Good 14 - luxury</td>
<td>0.22</td>
<td>0.39</td>
</tr>
<tr>
<td>With income tax, inequality aversion = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 1 - necessity</td>
<td>-0.01</td>
<td>0.40</td>
</tr>
<tr>
<td>Good 12 - luxury</td>
<td>0.57</td>
<td>0.79</td>
</tr>
</tbody>
</table>

First, in all cases the decrease in optimal tax rate exceeds one third of the evasion rate. With two necessities, the reduction in tax exceeds the evasion rate itself. Hence, as a conservative rule, it can be said that usually evasion leads to a fall in optimal tax rates exceeding half the rate of evasion, when dead-weight costs amount to a quarter or more of the revenue lost. Evidently, the reduction in tax rates is much larger for necessities than for luxuries. The approximate formula in (28a) indicates the opposite. With the utility ratio \( \bar{u}_{mi} \) in the negative term, one would expect the fall in tax rate to be larger for luxuries, given that with these goods \( \bar{u}_{mi} \) will be larger. The outcome from the iterative calculations suggests that with necessities the optimal solution tries to reduce evasion related dead-weight costs on low income households, by reducing the tax rate, which from (26) 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. Given the larger reduction in taxes on necessities, evasion reinforces the progressivity of optimal indirect taxation. It should be noted that my earlier evasion model, reported in Revesz (1997, 2014a), also revealed a highly progressive impact of evasion on optimal tax rates, provided it involves substantial dead-weight costs.

The fact that contrary to (28), the reduction in the tax rates of necessities is much larger than those of luxuries, suggests that the predicted rates using (28) are not particularly accurate. Indeed, comparison of actual and predicted rates in Table D.3 in appendix 3.4 reveals fairly large discrepancies. Since it appears that the optimisation process tries to minimise the dead-weight costs on low income taxpayers, it appears that an appropriate adjustment factor should be included, in addition to the two terms in (28). I found that an adjustment factor that works reasonably well is \( \frac{d_i}{\bar{u}_{mi}} \), when the benefits of evasion are spread among the consumers of product i in direct proportion to consumption. When the benefits are spread according to distribution weights, an adjustment factor \( -\frac{d_i u_i}{\bar{u}_{gi} \bar{u}_{mi}} \) is probably better. Following the incorporation of these adjustment factors, the approximate optimal tax formulas in (28a) and (28b) will become:
\[
\Delta t_i(e_i) \approx (1 - d_i)e_i - e_i u_{mi} - d_i / u_{mi}
\] (30a)
and
\[
\Delta t_i(e_i) \approx (1 - d_i)e_i \frac{u_{gi}}{u_{ii}} - e_i u_{mi} - d_i u_i / u_{mi}
\] (30b)

As shown in Table D.3 in appendix 3.4, the incorporation of these rather ad hoc adjustment factors tends to improve the accuracy of the predicted tax rates. But in some cases, this adjustment can lead to less accurate predictions than formula (28) without it. Perhaps, there is scope for improving the accuracy of these predictions, but I prefer to leave this subject to future research.

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. The experience with petrol and cigarette taxes suggests that tax rates in excess of 100% can be imposed on goods produced and marketed through large organisation, without encountering insurmountable evasion difficulties. In section 3.4.3 we noted that under current legal-institutional arrangements, separate commodity tax rates have to be defined for a small number of broad products groups. In light of the results from the present model, setting significantly higher tax rates for most non-necessities that are produced by large organisations, seems to be an advisable approach to increase consumption taxes, in line with the recommendations of the “growth friendly” taxation reports.

3.5.4 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. 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 expenditure on an externality generating good as \( Q_i \), then
\[
E_i = \mu_i Q_i
\] (31)
where \( \mu_i \) 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:
\[
U = \sum_h a_h w_h (P, W_h, \sum_i [y_{oh} + z_{ih} \mu_i Q_i])
\] (32)

---

62 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).
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 real output decreases or increases by $E_i$.

In appendix 3.2 eq. (A.27), we derive an approximate formula for the impact of a single externality on the optimal tax rate of the externality generating good. After taking out the similar $\bar{e}_i$ and $\bar{e}_i(u_0)$ 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.

$$\Delta t_i(\mu_i) \approx - \frac{\mu_i \sum_h z_{ih} \partial u_h / \partial y}{\theta_i / \partial y}$$

(33)

The term in the denominator is the average marginal utility of product $i$, that was defined in (10), and more explicitly in appendix 3.2 eq. (A.13). The term in the numerator is a weighted average of the social marginal utilities of externality 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 marginal utility of the demogrant in (A.2), the numerator in (33) becomes $\mu_i \sum_h 1/H \partial u_h / \partial y = \mu_i \partial u / \partial b$. From definition (10) this implies that in this case:

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

(34)

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, its 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 (33) 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 (33) 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 to (33) for pollution externalities.63

---

63 Sandmo (1975) uses similar specifications to the present model, including no income tax and a single externality tax. Sandmo did not claim that the optimal pollution tax rate will usually exceed the Pigovian rate in the absence of income tax, because he did not test numerically his analytical model.
Having established an extension of the modified inverse elasticity rule for externalities, we can take a look at some numerical results. Table 7 presents results with constant proportional Pigovian 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 (33), the departure from the Pigovian rates (plus or minus 20%) is larger the more the externality affects low income persons (decreasing \( z_{ih} \)). Generally luxury goods do not show significantly larger deviations from the Pigovian rate than necessities, contrary to what is predicted from (34). It is interesting to note in Table 7 that without progressive income tax, actual changes in tax rates tend to be well above Pigovian rates, but in line with changes predicted from the extensions to the modified inverse elasticity formulas in (33) and (34). Recall that zero income tax combined with a demogrant is a particular form of linear income tax. When progressive income taxation is included in the model, the results are closer to Pigovian rates. These findings suggest that the presence or absence of other distributional instruments (such as progressive income tax), has a strong bearing on the impact of externalities on optimal tax rates.

Turning to the highly publicised subject of environmental taxes, it can be said that the material presented in this chapter supports higher environmental taxes. One reason is the finding that tax increases due to negative externalities will exceed the Pigovian tax rates with linear income tax, and to a lesser extent, will also tend to exceed them in the presence of progressive income tax. The other reason is connected with the fact that most fossil fuels are marketed

\[ \text{Table 7} \]

<table>
<thead>
<tr>
<th>Product no. and type</th>
<th>Pigovian rate</th>
<th>Marginal utility ratio*</th>
<th>Type of weights</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>All equal</td>
<td>Increasing</td>
</tr>
<tr>
<td>With progressive income tax – inequality aversion = 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 5 - necessity</td>
<td>-0.2</td>
<td>1.08</td>
<td>0.21</td>
</tr>
<tr>
<td>Good 17 - luxury</td>
<td>+0.2</td>
<td>2.26</td>
<td>-0.17</td>
</tr>
<tr>
<td>Two composite goods with income tax – inequality aversion = 0.3**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 5 - necessity</td>
<td>-0.2</td>
<td>0.91</td>
<td>0.16</td>
</tr>
<tr>
<td>Good 17 - luxury</td>
<td>+0.2</td>
<td>1.71</td>
<td>-0.13</td>
</tr>
<tr>
<td>Without income tax – inequality aversion = 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 4 - necessity</td>
<td>-0.2</td>
<td>1.43</td>
<td>0.46</td>
</tr>
<tr>
<td>Good 12 - luxury</td>
<td>+0.2</td>
<td>2.44</td>
<td>-0.47</td>
</tr>
<tr>
<td>Without income tax – inequality aversion = 0.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Good 4 - necessity</td>
<td>-0.2</td>
<td>1.15</td>
<td>0.27</td>
</tr>
<tr>
<td>Good 12 - luxury</td>
<td>+0.2</td>
<td>1.27</td>
<td>-0.29</td>
</tr>
</tbody>
</table>

* Marginal utility ratio in the situation without externalities
** In the two composite goods model the impact of externalities represents the difference between the tax rate on the externality generating good and the standard tax rates on other commodities belonging to the same composite good.

Having established an extension of the modified inverse elasticity rule for externalities, we can take a look at some numerical results. Table 7 presents results with constant proportional Pigovian 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 (33), the departure from the Pigovian rates (plus or minus 20%) is larger the more the externality affects low income persons (decreasing \( z_{ih} \)). Generally luxury goods do not show significantly larger deviations from the Pigovian rate than necessities, contrary to what is predicted from (34). It is interesting to note in Table 7 that without progressive income tax, actual changes in tax rates tend to be well above Pigovian rates, but in line with changes predicted from the extensions to the modified inverse elasticity formulas in (33) and (34). Recall that zero income tax combined with a demogrant is a particular form of linear income tax. When progressive income taxation is included in the model, the results are closer to Pigovian rates. These findings suggest that the presence or absence of other distributional instruments (such as progressive income tax), has a strong bearing on the impact of externalities on optimal tax rates.

Turning to the highly publicised subject of environmental taxes, it can be said that the material presented in this chapter supports higher environmental taxes. One reason is the finding that tax increases due to negative externalities will exceed the Pigovian tax rates with linear income tax, and to a lesser extent, will also tend to exceed them in the presence of progressive income tax. 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.

3.5.5 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. This implies a difference between private and social preferences (Kanbur, 2006). 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. Related products are sometimes referred to in the literature as merit (or demerit) goods. Some interesting papers on this subject are presented by Besley (1988) and Pestieau et al. (2012).

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 per unit expenditure is denoted \( \gamma_i \), therefore total merit goods 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 with a merit good will be:

\[
U = \sum_h a_h u_h (P, W_h, \sum_i [y_{oh} + Q_{ih} \gamma_i])
\]

Comparing this expression with the social welfare function involving externalities defined in (32), we notice that the only difference is that the \( \sum_i [y_{oh} + z_{ih} \mu_i Q_i] \) term of externalities, has been replaced by the \( \sum_i [y_{oh} + Q_{ih} \gamma_i] \) term here. Therefore, there is no need to develop another extension to the modified inverse elasticity rule. Suffice is to replace the externality term by the merit good term. As with externalities, we focus is on a single good – \( Q_i \). From (33) it follows that in this case:

\[
\Delta t_i(\gamma_i) \approx -\frac{\gamma_i \sum_h q_{ih} \partial u_h/\partial y}{\partial u_i/\partial y} = -\frac{\gamma_i \partial u_i/\partial y}{\partial u_i/\partial y} = -\gamma_i
\]

The simplification follows from the definition of the average marginal utility of product \( i \) in appendix 3.2 eq. (A. 13).

Table 8

<table>
<thead>
<tr>
<th>Product No. and type</th>
<th>Imputed ( \gamma_i ) with paternalistic objective</th>
<th>without such objective</th>
<th>difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good 6 - necessity</td>
<td>-0.2</td>
<td>0.17</td>
<td>0.42</td>
</tr>
<tr>
<td>Good 16 - luxury</td>
<td>+0.2</td>
<td>1.48</td>
<td>1.28</td>
</tr>
<tr>
<td>Good 2 - necessity</td>
<td>+0.2</td>
<td>0.55</td>
<td>0.30</td>
</tr>
<tr>
<td>Good 14 - luxury</td>
<td>-0.2</td>
<td>0.14</td>
<td>0.38</td>
</tr>
</tbody>
</table>

With income tax – inequality aversion = 1

Without income tax – inequality aversion = 0.3
Having obtained a simple approximation for the effect of $\gamma_l$ on the optimal tax rate, we can take a look at some numerical examples in table 8. Generally, the changes in optimal tax rates from the iterative calculations are in line with what is predicted from (36).

### 3.5.6 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 between commodities and leisure, it is of some interest to examine the pattern of optimal taxation controversy. All the models favouring uniform taxation assume weakly separable rates from the iterative calculations are in line with what is predicted from (36).

We can use $\varepsilon_i$ in a modified inverse elasticity calculation to estimate its approximate effect on optimal tax rates. The detailed derivation is presented in appendix 3.5. The end result is the following approximation from (E.4):

$$\Delta t_i (\ell) \approx - \frac{\partial \varepsilon_i}{\partial q_i} \bar{u}_{mi} \bar{W}_i T_i' = \frac{-\varepsilon_i}{\bar{q}_i} \bar{u}_{mi} \bar{W}_i T_i'$$

where

$$\frac{\partial \varepsilon_i}{\partial q_i} = \frac{\partial \varepsilon_i/\partial q_i/\partial \ell_l}{\partial \ell_l}$$

$\varepsilon_i$ is the $q_i$ weighted average value of $\varepsilon_i$ over $H$ taxpayers, $\bar{W}_i$ is the corresponding average wage rate, $T_i'$ is the average marginal indirect tax rate of the consumers of product $i$. It is defined in (A.5b) in appendix 3.2. $\bar{u}_{mi}$ is the marginal utility ratio of product $i$, defined in (A.12). The approximation in (40) suggests that the tax rate on $i$ will increase (or decrease) depending by how much tax revenue decreased (or increased) as a result of a change in

---

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

64
labour supply. Another point to notice in (40) 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

<table>
<thead>
<tr>
<th>good no.</th>
<th>$\frac{\partial T_i}{\partial q_{i}}$</th>
<th>$\bar{W}_i$</th>
<th>$\bar{T}_i$</th>
<th>$\bar{u}_{mi}$</th>
<th>original model tax</th>
<th>modified model tax</th>
<th>actual tax difference</th>
<th>predicted tax difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.03</td>
<td>17.6</td>
<td>0.55</td>
<td>1.22</td>
<td>0.98</td>
<td>0.20</td>
<td>-0.78</td>
<td>-0.35</td>
</tr>
<tr>
<td>2</td>
<td>18.3</td>
<td>0.55</td>
<td>1.28</td>
<td>0.85</td>
<td>1.02</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>13.2</td>
<td>0.50</td>
<td>0.95</td>
<td>0.22</td>
<td>0.44</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>-0.02</td>
<td>20.1</td>
<td>0.56</td>
<td>1.41</td>
<td>0.98</td>
<td>1.83</td>
<td>0.85</td>
<td>0.32</td>
</tr>
<tr>
<td>5</td>
<td>17.9</td>
<td>0.53</td>
<td>1.16</td>
<td>0.48</td>
<td>0.71</td>
<td>0.23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>19.7</td>
<td>0.56</td>
<td>1.37</td>
<td>0.98</td>
<td>1.15</td>
<td>0.17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>0.01</td>
<td>15.9</td>
<td>0.53</td>
<td>1.11</td>
<td>0.45</td>
<td>0.51</td>
<td>0.06</td>
<td>-0.09</td>
</tr>
<tr>
<td>8</td>
<td>17.8</td>
<td>0.54</td>
<td>1.23</td>
<td>0.68</td>
<td>0.87</td>
<td>0.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>-0.04</td>
<td>15.5</td>
<td>0.52</td>
<td>1.09</td>
<td>0.43</td>
<td>0.70</td>
<td>0.27</td>
<td>0.17</td>
</tr>
<tr>
<td>10</td>
<td>0.05</td>
<td>27.8</td>
<td>0.61</td>
<td>2.26</td>
<td>1.49</td>
<td>0.86</td>
<td>-0.62</td>
<td>-1.90</td>
</tr>
<tr>
<td>11</td>
<td>25.9</td>
<td>0.61</td>
<td>2.12</td>
<td>2.68</td>
<td>2.52</td>
<td>-0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>27.6</td>
<td>0.61</td>
<td>2.24</td>
<td>1.62</td>
<td>1.77</td>
<td>0.15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>-0.02</td>
<td>28.3</td>
<td>0.61</td>
<td>2.30</td>
<td>1.55</td>
<td>2.83</td>
<td>1.28</td>
<td>0.79</td>
</tr>
<tr>
<td>14</td>
<td>28.81</td>
<td>0.61</td>
<td>2.34</td>
<td>1.32</td>
<td>1.53</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>26.5</td>
<td>0.61</td>
<td>2.17</td>
<td>2.12</td>
<td>2.14</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>26.2</td>
<td>0.61</td>
<td>2.15</td>
<td>2.33</td>
<td>2.29</td>
<td>-0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>28.9</td>
<td>0.61</td>
<td>2.35</td>
<td>1.31</td>
<td>1.52</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>-0.14</td>
<td>27.8</td>
<td>0.61</td>
<td>2.26</td>
<td>1.70</td>
<td>3.97</td>
<td>2.27</td>
<td>4.34</td>
</tr>
</tbody>
</table>

Average tax rate on the 9 necessities 0.70 0.67
Average tax rate on the 9 luxuries 1.83 1.93
% change in welfare terms compared to uniform 3.5 15.0
% change in output compared to uniform solution 1.7 19.9

* The complementarity parameter on product 9 applies to the preferences of the 8 lower W taxpayers. The one on product 18 applies to the 7 higher ability taxpayers.

Apparently, 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 extension to the modified inverse elasticity rule is based on the assumption that the introduction of $\epsilon_i$ has no first order effect on the incomes of taxpayers. Only its impact on tax revenue has first order effect (see appendix 3.5). This perspective on the role of $\epsilon_i$ 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.
In an analytical study, Jacobs and Boadway (2014) examine non-separable utility in a many-person model, with commodity taxes combined with a Mirrleesian optimal non-linear income tax function. They reached the same conclusions as Corlett and Hague (1953) in the context of this redistributive model.65

Table 9 compares actual tax rates from a numerical study, with the rates predicted using the modified inverse elasticity rule in (40). This numerical study is based on an extended form of LES, which incorporates leisure complements and substitutes. It is explained in detail in appendix 3.5. The original model we compare with is the one reported in Table 1, with the inequality aversion rate set to one. Notice in Table 9 that larger changes involving luxuries, has led to 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 $\frac{\partial \tilde{y}}{\partial q_t}$ parameters are not based on econometric estimates but were arbitrarily chosen, and the ad hoc nature of extended LES used in these calculations (see appendix 3.5), perhaps one should observe with caution these striking results.

The results from Table 10 indicate that the large improvements over the uniform solution arise because of the high inequality aversion rate chosen. Table 10 shows that the utility gains over the uniform solution become much smaller when the inequality aversion rate is reduced from one to 0.3. The explanation is that with lower inequality aversion, there is less utility gained by boosting labour supply in order to generate more revenue for redistribution. The scenarios in Table 10 are based on the same parameters reported in Table 9.

<table>
<thead>
<tr>
<th>Inequality aversion rate</th>
<th>Original model</th>
<th>Extended model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% change welfare over uniform</td>
<td>% change output over uniform</td>
</tr>
<tr>
<td>Without progr. income tax</td>
<td>0.3</td>
<td>0.2</td>
</tr>
<tr>
<td>With progr. income tax</td>
<td>1</td>
<td>2.8</td>
</tr>
<tr>
<td>With progr. income tax</td>
<td>0.3</td>
<td>0.4</td>
</tr>
</tbody>
</table>

65 West and Williams (2007) and Parry and West (2009) present estimates of leisure complementarities, and their effect on optimal tax rates in a partial equilibrium setting.
3.6 Summary and qualifications

3.6.1 The main results

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 persists even when the inequality aversion of society is low. Subject to lower dispersion and progressivity, this conclusion holds true when exogenously given income tax schedules are incorporated into the model. Moreover, the gains in output and welfare of the differentiated over the uniform solution can be quite substantial. These gains can sometimes exceed 3% of total output, even in a simple 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 negative and nearly uniform, when inequality aversion is low. Yet, while exceptions are possible, the large majority of numerical results in purely redistributive models point in favour of differentiated and progressive indirect 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, given zero or linear income tax, externality generating goods should be taxed or subsidised well above Pigovian rates.

3.6.2 Some qualifications

A number of possible objections can be raised against the present model, 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 indicated in chapter 2, 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

---

66 It is likely that factors with large variations between products, such as externalities, leisure complementarities or evasion propensities, will have a larger effect on tax rate differentiation than factors with smaller differences between products, such as administrative and compliance costs.
theorem by Laroque (2005) and Kaplow (2006) has been analysed in section 3.4.5 and its limitations were noted.

Yet another objection can be raised about using arbitrary numbers in the simulations rather than empirical estimates. Given that much of the discussion is based on analytical foundations, this does not appear to be a major problem. Analytical studies from the 1970s and 80s have already indicated that with zero or linear income tax, optimal commodity tax rates will be differentiated and progressive (see sections 2.2.1 and 3.4.1). Other factors studied in this chapter were analysed using the modified inverse elasticity rule. Hence, the broad conclusions presented here are analytically based, and do not depend on the choice of parameter values. The numerical examples illustrate and substantiate analytical arguments. Hopefully, future numerical research on this subject will rely more on empirical estimates.

Another possible objection relates to the non-linear Engel curves in the present model. The segmentation of LES utility has led to a sharp distinction between necessities and luxuries, which could be considered unrealistic and might exaggerate the benefits of progressive indirect taxation. This is probably a significant weakness in the present model. It might be advisable to use in future research utility functions other than segmented LES, or divide LES into more than two segments.

3.6.3 Policy issues

When examining possible policy implications from the present study, I shall focus on European Union (EU) VAT systems, since this is an area that I am a bit familiar with. Currently standard VAT rates in EU countries range between 15% and 27%. A minority of products (mainly necessities) are subject to zero or reduced VAT rates. These concessions are more significant in the UK, France, Italy and Spain. We shall refer to these systems as “progressive” VAT systems.

Considering the legal-institutional requirement to define tax rates for broad product groups, it appears that there is little room to provide further tax relief on necessities beyond what is available in “progressive” VAT systems, apart from setting all concessional rates to zero. In contrast to necessities in “progressive” VAT systems, there appears to be plenty of scope for reform in regard to non-necessities. As noted earlier, tax rates in excess of 100% have been applied for a long time on petrol and cigarettes. Such high tax rates are likely to be administratively manageable also in respect to other non-essential goods and services produced by large organisations. Our numerical results suggest that tax rates on non-necessities in excess of 50% can be justified even under moderate inequality aversion rates. For products that generate negative externalities, taxes should be raised even higher. If the objective is to increase revenues from consumption taxes, as recommended in the “growth friendly” taxation studies, then there seems to be plenty of room to increase VAT rates on non-necessities produced and/or marketed by large organisations. For that purpose, it might be advisable to introduce a few

67 In fact, the utility-demand parameters for the low wage group (see Table C.1) are based on econometric estimates by Deaton, and Muellbauer (1980). For further details see Revesz (1997). The parameters chosen for the higher wage group are arbitrary.

68 As explained in section 3.4.4, for the purpose of providing support to those at the bottom end of the scale, the targeting and monitoring of welfare payments and in-kind transfers is probably a more critical economic issue than reduced taxation of necessities.
separate VAT rates for non-necessities, instead of one standard rate, and assign tax rates to product groups taking into account income elasticities, evasion propensities, administrative and compliance costs, paternalistic concerns, leisure complementarities and legal definitional requirements. Generally speaking, making indirect taxation more progressive, can offset the erosion of equity goals due to income tax cuts, which have been recently implemented or have been proposed in many countries.

In regard to the taxation of housing, it is possible to introduce progressive taxation in this area through the existing system of property valuations. But given that property taxes are usually levied by local authorities, who have little interest in distributional issues, progressive property taxation probably will have to be implemented by central authorities. In summary, looking at current EU indirect taxation systems, there appears to be considerable scope for welfare improving and evasion reducing reforms. I prefer not to speculate on the political feasibility of such reforms.

All in all, the present 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 3.1**

**User’s guide to the computer program titled ´indirect-tax8´**

This program carries out optimisation of commodity tax rates under a broad range of specifications. It yielded the numerical results discussed in this book. 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 20 years ago and was reported in Revesz (1997). The current version of the program is much broader than the original.

**Program installation and running**

The program is located in my website [https://johnlreveszcom.netfirms.com/index.html](https://johnlreveszcom.netfirms.com/index.html).

It is written in a programming language called QBasic, which two decades ago was part of the tool kit 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/](http://www.qb64.net/). Another QBasic compiler is available for US $60 from: [http://www.libertybasic.com/index.html](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-tax8 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. This menu provides access to various options such as type of utility function, inequality aversion rate, the inclusion of income tax, tax evasion, administration, compliance, externalities, paternalistic concerns and leisure complements-substitutes. The prompts also present a few parameter choices, as well as choices of scenarios in respect to tax evasion, externalities and other consumption related factors.

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. For further calculations and analysis, the output file can be converted into an Excel spreadsheet by pasting tax1.txt into an Excel file, choosing the Data option and then performing a text to column conversion with space as the delimiter.

In case you encounter any problem when responding to prompts, the quickest way to overcome it, is to exit and start again. Click on the cross at the upper right hand corner of the QBasic screen. After the screen disappears, start running the program again.

Finally, all the library files associated with the compiler (qb64) 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 altered. Above the DATA lines are text lines (marked at the left by 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-demand 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-substitutes. Some of these parameters will be discussed later. Others are self-explanatory.

Usually there are a number of DATA lines for each item. Those marked with an asterisk at the beginning of the line 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. Always make sure that there is only one DATA line without an asterisk for each variable. Note, every time you change DATA lines on the indirect-tax8 Word file, you have to copy and paste the file again into the QBasic screen. However, you can also perform DATA editing on the QBasic screen itself.
Utility and demand parameters

The current version of the program enables the user to specify utility-demand parameters according to his/her choice for the two wage groups in the model. The relevant DATA lines for 9 goods and 18 goods utility-demand appear close to the beginning of the source code. The leisure parameters are defined under them through equations with numbers on the right hand side. The program checks that the betas of commodities and leisure add up to one. If this condition is not satisfied a warning will appear at the very start of the printout. To ensure the global quasi-concavity of LES utility, all betas must be positive.

The original (Revesz 1997) utility and demand parameters used in the program can be applied also in the current version. These appear in two DATA lines for the arrays beta0 and gamma0, with the warning that these data should not be altered. These specifications are referred to in Revesz (2014a) as the “bi-polar” scenario. By varying the split factor (through a prompt) between 0.01 and 0.99, a broad range of hypothetical demand systems can be generated, ranging from almost linear (with a split factor of 0.01) to highly non-linear (with a split factor of 0.99). Using a split factor of 0.99, the results reported in Revesz (1997) can be replicated fairly closely. You obtain bi-polar specifications if you type in a number between zero and one, in response to the prompt asking about the split factor that will define the bi-polar parameters. Typing in any number above one, will direct the program to use the parameters specified separately for the two wage groups, as discussed earlier.

The income tax schedule

The program includes an income tax schedule. This is a piecewise linear schedule defined by five marginal tax rates over five income intervals. The marginal tax rates are specified (through a DATA line) for 5 groups with 3 taxpayers in each. The corresponding initial income intervals are presented in the initial income parameters table appearing close to the beginning of the printout. Given that taxable income is a dependent variable, the program will change the marginal tax rate of a taxpayer, if this is warranted due to changing labour supply. Taxable income is defined as gross income ($W$) minus the portion of tax evaded by the taxpayer, which is specified by another DATA line. The five bracket piecewise linear income tax schedule can approximate a wide range of shapes. You can access the income tax option by replying with “yes” to the relevant prompt.

Inequality aversion

The original program included two social welfare functions to describe the inequality aversion of society, one is LES utility (a logarithmic function) the other is maximin. By definition, maximin leads to high optimal tax rates. Even LES leads to fairly high tax rates. To overcome this problem, an inequality aversion rate ($z$), ranging between 0 and 1 was introduced, as explained in section C4 in appendix 3.3. The user can choose this rate in the second prompt. Sometimes when $z$ is low, the search for the optimum does not converge properly. In these cases an error message is issued by the program. When this happens, then changing parameters to increase taxation requirements, such as higher inequality aversion or more tax evasion or expenditure on public goods, could help to overcome the problem.
**Tax evasion and administration**

The program includes tax evasion on certain commodities as a fixed portion of tax, with an associated fixed percentage of dead-weight losses. It also includes administration and compliance costs as a fixed percentage of the tax collected. According to the definitions used, the dead-weight costs of administration are paid by the government while those of compliance are borne by consumers. These items are defined separately for commodity and income taxation. The relevant parameters appear in DATA lines in the source code.

The prompt asking whether to include in the model tax evasion and administrative costs covers a number of sub-items. In regard to most of these sub-items no further prompts will appear. On the commodities side these include compliance costs by consumers, administrative costs by government and tax evasion involving effective price reductions. In regard to compliance, administration and evasion related to income taxation, no further sub-options will be presented. To use any of these options, you only have to make sure that the relevant DATA lines are filled in according to your choice. The inclusion of administration-evasion parameters for both commodity and income taxation, opens the way for exploring the possibilities for achieving a better tax mix depending on these factors.

The program contains two principal indirect tax evasion scenarios. (i) The income transfer scenario is described in this book. (ii) The effective price reduction scenario is described in Revesz (2014a). In regard to the income transfer scenario, the program also asks whether the distribution of evasion benefits should be based on the consumption of the good concerned, or should it be based on fixed pre-determined weights that add up to one.

Another prompt relates to the scenario of commodity tax rate dependent evasion, which only applies above the average tax rate (see section 5 in Revesz (2014a)). The relevant parameters are specified in a separate DATA line.

**Externalities and other consumption related effects**

This prompt covers three sub-items. These are externalities, paternalistic concerns and leisure substitutes and complementarities. We shall discuss these items one by one.

**Externalities**

Externalities appear as a fixed percentage of sales, and represent imputed costs or benefits related to the total consumption of certain goods. Negative externalities should be entered into the respective DATA lines as negative numbers and vice versa for positive externalities. In addition, data on the distribution weights of externality recipients should be included. If the externality weights do not add up to 15 (the number of taxpayers), the program issues an error message.

**Paternalistic concerns**

The paternalistic option appears in the menu as a sub-option. Imputed benefits related to paternalistic concerns (merit goods), should be entered as positive percentages of net sales and imputed costs (demerit goods) as negative percentages.
Leisure complement and substitutes

Leisure complementarity appears in the menu as a sub-option, following the prompt that asks about the inclusion of “externalities or other consumption dependent costs or benefits”. There is one DATA line related to this item, referring to the $k_i$ coefficients explained in appendix 3.5. Leisure substitutes should be entered as positive numbers and vice versa for leisure complements. The program ensures that the zero homogeneity condition $\sum_i k_i q_{10} = 0$ (see appendix 3.5) is satisfied. In the higher $W$ taxpayers’ preference maps, $k_{18}$ is used to ensure the zero sum condition. For the lower $W$ taxpayers it is $k_9$. Therefore, in the relevant DATA line, the $k_i$ for commodities 9 and 18 should be set to zero. When you use the leisure substitution-complementarity option, please make sure that not all the $k_i$ are set to zero.

Capital income

The program includes capital income. It is represented by pre-tax lump-sum income. Under competitive conditions, it is reasonable to assume that non-work related lump-sum income is equivalent to capital income. The present static model is not really suitable to examine capital taxation issues. Nonetheless, the inclusion of capital income could be used to examine its effect on optimal commodity tax rates. There is an option to tax capital income at a higher or lower rate than labour income. To describe the common situation of identical treatment in income tax, please type in one for the capital income tax multiplier in response to the prompt.

Targeted support grants

If you answer with ‘yes’ to this prompt, the program will add the values in the respective DATA line to the lump-sum incomes of the lowest $W$ taxpayers. These grants are supposed to represent lump-sum support payments to the needy that are given exogenously, on the basis of considerations outside the model, unlike support provided through the demogrant.

Fixed expenditure on public goods

You can specify expenditure on public goods as a percentage of total output in the pre-tax situation, when the demogrant is zero. Following the introduction of taxation the demogrant will increase, and due to its disincentive effect on labour supply, total output will decrease. Hence the expenditure on public goods will represent a higher percentage of total output in the post-tax situation than the percentage you have specified.

The two composite goods model

This option appears in the latest version of the program. Following an affirmative answer to the relevant prompt, the program will go the DATA lines listed under “bypass3”. Given the same beta and gamma parameters for all the necessities and another set of identical parameters for all luxuries, the program will effectively treat them as two goods instead of 18 goods.

Sub-optimal demogrant

The optimal demogrant option is included in the latest version of the program. After typing in ‘yes’ to this prompt, the program will ask what percentage should be the fixed demogrant from pre-tax average income. You specify a percentage according to your choice,
including possibly also zero. The percentage multiplied by total pre-tax income will replace Hb in constraint (2a). Otherwise, the iterative optimisation process will follow the same procedure as in other cases.

**The equivalent income tax calculations**

These calculations show the results of replacing all commodity taxes by an equal revenue generating income tax at each W level. It is explained in detail in Revesz (1997). Occasionally there have been some difficulties with these calculations, particularly with maximin and/or when a non-linear income tax schedule is present. The source of these difficulties lies in the fact that with LES utility, the consumption and labour supply values are given by a system of equations (see section 3.3). If due to equivalent marginal income tax rates in excess of 70%, some taxpayer(s) choose to supply negative work according to the leisure demand equation, it seems logical to put their labour supply to zero, rather than to leave it as a negative number. However, when one of the results is altered from the calculated outcome, the system of equations will no longer yield consistent results.

More explicitly, the following accounting identity between income and consumption will be violated. Consumption expenditure (at factor prices) = Gross income – taxes + demogrant – dead-weight costs. When this identity is violated, the results will no longer be internally consistent. When this happens, I decided not to show the equivalent income tax results, rather than to present unreliable numbers.

**The utility adjustment constant**

Because segmented LES consists of two separate cardinal utility functions for two population groups, it does happen sometimes that the utility level of the ninth taxpayer, belonging to the high W group, is smaller than the utility of the eighth taxpayer, who belongs to the lower income group and has a lower W and smaller income. In other cases there is a sudden jump in utility between the eighth and the ninth taxpayers. To overcome such dip or jump, the program calculates a “smoothing” constant that is added to the utilities of all seven members of the higher W group. This constant does not affect in any way the calculations, and prevents the appearance of an illogical result. Given that the same constant is added to the utilities of all members of the high W group, in both the uniform and non-uniform tax calculations, the utility outcomes from these calculations are comparable. But the problem with this utility constant is that it changes with each scenario. In order to enable the user to make utility comparisons between different scenarios, the printouts show the constant added to the utilities of the higher W group. By taking away this constant from the utilities of the seven higher wage earners, the user can obtain utility values that are more comparable between different scenarios.
Appendix 3.2
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 simple distributional model and has been developed further to accommodate additional factors.

In the model without income tax (analysed in section 3.4.1), 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

$$\sum_h \frac{\partial u_h}{\partial t_i} = - \sum_h q_{ih} \frac{\partial u_h}{\partial y}$$  \hspace{1cm} (A.1)

where we sum up over $H$ taxpayers.\(^{69}\)

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

$$\frac{\partial U}{\partial b} = \frac{\sum_h q_{yh}}{H}$$  \hspace{1cm} (A.2)

The amount transferred depends on the revenue generated by increased $t_i$. In order not to complicate the analysis, we ignore for a moment changes in labour supply. In this situation, deriving government revenue ($R = \sum t_i q_i$) with respect to $t_i$ yields:

$$\frac{\partial R}{\partial t_i} = q_i + \sum_j t_j \frac{\partial q_j}{\partial t_i}$$  \hspace{1cm} (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 $t_i \frac{\partial q_i}{\partial t_i}$. When all tax rates are the same ($t_i$), and provided labour supply does not change, then

$$\sum_j t_j \frac{\partial q_j}{\partial t_i} = t \sum_j (1 + t) \frac{\partial q_j}{\partial p_j} = t \sum_j p_j \frac{\partial q_j}{\partial p_j} = 0$$  \hspace{1cm} where $q_i$ represents total demand by $H$ taxpayers. These cases suggest that the total revenue derivative may be approximated by the following expression:

$$\frac{\partial R}{\partial t_i} = q_i + \sum_j t_j \frac{\partial q_j}{\partial t_i} = q_i + \frac{\partial q_i}{\partial t_i} (t_i - t_{iA})$$  \hspace{1cm} (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. Finally, we take into account the effect of changing labour supply on tax revenue. Multiplying the cross derivative of labour supply with respect to $p_i$ by $W$ yields the change in gross income due to the tax change. Symbolically: $\frac{\partial m}{\partial p_i} = \frac{\partial t}{\partial p_i} W$. Multiplying the change in income by the average portion of

---

\(^{69}\) Note, here utility is derived with respect to lump-sum income ($y$) rather than income ($m = W + y$), but with weakly separable utility the two derivatives are the same (see Revesz (1986)).
commodity taxes from the last dollar expenditure of the consumers of product i (in other words, the combined marginal indirect tax rate, denoted $\bar{T}_{il}'$), we obtain the change in revenue due to labour supply:

$$\frac{\partial R(\ell)}{\partial t_i} = \frac{\partial \ell}{\partial p_{li}} W_i \bar{T}_{il}'$$  \hspace{1cm} (A.5a)

where

$$\bar{T}_{il}' = \sum_j t_j \frac{\partial q_{ji}}{\partial y} \left(1 + \sum_j t_j \frac{\partial q_{ji}}{\partial y} \right)$$  \hspace{1cm} (A.5b)

Here $q_{ji}$ refers to the quantity consumed from $q_j$ by the consumers of $q_i$. This definition is needed in order to express $\bar{T}_{il}'$ as a fraction of income, just like $\bar{T}_{mi}$. Combining (A.2), (A.3) and (A.5), the net effect on social utility of giving transfers through the demogrant following a change in $t_i$ is defined by the following expression:

$$\frac{\partial u}{\partial b} \frac{\partial R}{\partial t_i} \approx \frac{\partial u}{\partial b} \sum_h (q_{ih} + \frac{\partial q_{ih}}{\partial p_{li}} (t_i - t_{iA})) + \frac{\partial \ell}{\partial p_{li}} W_h \bar{T}_{il}'$$  \hspace{1cm} (A.6)

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:

$$D = -\Delta t_i \Delta q_i / 2 \approx -\Delta t_i \frac{\partial q_{i(u_0)}}{\partial t_i} \frac{\Delta t_i}{2} \approx -t_i \frac{\partial q_{i(u_0)}}{\partial t_i} \frac{\Delta t_i}{2} \approx -t_i^2 \frac{\partial q_{i(u_0)}}{\partial p_{li}} / 2$$  \hspace{1cm} (A.7)

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.7), $\Delta t_i = t_i$. The excess burden loss is borne directly by the taxpayer. Deriving (A.7) with respect to $t_i$ and then multiplying by the consumer’s marginal utility of income term and summing over $h$ taxpayers we obtain:

$$\frac{\partial u}{\partial t_i} \approx -\sum_h \frac{\partial u_h}{\partial y} t_i \frac{\partial q_{i(u_0)}}{\partial p_{li}}$$  \hspace{1cm} (A.8)

At the optimum, the three utility effects, represented by (A.1), (A.6) and (A.8) must add up to zero. Summing up while converting derivatives into elasticities ($\varepsilon$) we obtain:

$$0 \approx -\sum_h q_{ih} \frac{\partial u_h}{\partial y} +$$

$$+ \frac{\partial u}{\partial b} \left[ \sum_h (q_{ih} + \frac{q_{ih}}{p_i} (t_i - t_{iA})) + \varepsilon_{ih}(\ell) m_{ih} \right] + \frac{\partial \ell}{\partial p_{li}} T_{il}' + \sum_h \frac{\partial u_h}{\partial y} t_i \frac{q_{ih} \varepsilon_{ih}(u_0)}{p_i}$$  \hspace{1cm} (A.9)

Multiplying all the terms in (A.9) by $p_i / \sum_h q_{ih} = p_i / Q_i$ we arrive at:

$$0 \approx - \frac{\partial u_i}{\partial y} p_i + \frac{\partial u_i}{\partial b} \left[ p_i + \bar{\varepsilon}_i (t_i - t_{iA}) + \bar{\varepsilon}_i (\ell) \bar{T}_{i}' + \frac{m_{i}}{Q_i} - t_i \frac{\partial u_i}{\partial y} \bar{\varepsilon}_i (u_0) \right]$$  \hspace{1cm} (A.10)

where $\frac{\partial u_i}{\partial y}, \bar{\varepsilon}_i, \bar{\varepsilon}_i (u_0), \bar{\varepsilon}_i (\ell), \bar{T}_{i}'$ are all weighted averages, with the weights given by the $q_{ih}$ purchases of individual consumers. After dividing all terms by $- \frac{\partial u_i}{\partial y}$ and replacing $p_i$ by 1 + $t_i$ and $\varepsilon_i (\ell) \bar{T}_{i}' + \frac{m_{i}}{Q_i}$ by $\bar{\varepsilon}_i (u_0) \bar{T}_{i}'$ we arrive at the following elasticities-based formula:

$$t_i \approx \frac{1 + \bar{u}_{mi} (1 + t_i + \bar{\varepsilon}_i (t_i - t_{iA}) + \bar{\varepsilon}_i \bar{T}_{i}')}{\bar{\varepsilon}_i (u_0)}$$  \hspace{1cm} (A.11)

The term

$$\bar{u}_{mi} = \frac{\partial u / \partial b}{\partial u_i / \partial y}$$  \hspace{1cm} (A.12)

is called the marginal utility ratio of product i. $\partial u / \partial b$ has been defined in (A.2). $\partial u_i / \partial y$ is the average marginal utility of the consumers of product i. It is defined as:
\[ u_i = \frac{\partial u_i}{\partial y} = \frac{\sum_h \partial u_h}{\sum_h u_h} q_{ih} \]  

(A.13)

The approximate formula in (A.11) is referred to as the modified inverse elasticity rule. A similar approximation presented in Sandmo (1975) eq. (24) appears in terms of the notations used here as:

\[ t_i \approx -\frac{1 - \frac{1}{\bar{u}_{mi}}}{\bar{e}_i} = \frac{1 - \bar{u}_{mi}}{\bar{u}_{mi} \bar{e}_i} \]  

(A.14)

Apparently, the biggest difference between Sandmo’s formula and the one developed here, is the absence of the \( \bar{e}_i (t_i - t_{iA}) \) and \( \bar{e}_{i\ell} \bar{T}_{i\ell} \) terms in Sandmo’s approximation.

When income tax is included, the modified inverse elasticity rule is changed. It will be:

\[ t_i \approx \frac{1 + t_i - \bar{u}_{mi} \left( 1 + t_i + \bar{e}_i (t_i - t_{iA}) + \bar{e}_{i\ell} (\bar{T}_{mi} + \bar{T}_{i\ell}) \right)}{\bar{e}_i (u_0)} - \bar{T}_i (1 - \gamma) \]  

(A.15)

where \( \bar{T}_{mi} \) is the average marginal income tax rate of the consumers of product \( i \), \( \bar{T}_i \) is the corresponding average income tax rate. \( \gamma \) is a general adjustment factor. It ensures that total tax revenue from (A.11), with the \( \bar{T}_{mi} \) term added, will equal to the combined total direct and indirect tax revenue from the iterations when income tax is included. It follows from this total revenue-equalising definition of \( \gamma \) that the average commodity tax rate from the iterations and the average predicted from (A.15) will be the same. \( \gamma \) is usually below plus or minus 0.15, indicating that the optimal combined tax rate \( \sum_i (t_i + \bar{T}_i) \) on the consumer group of product \( i \) in the presence of income tax, is not much different from the optimal combined tax rate without it (see appendix 2.1).

Given that (A.11) and (A.15) are used in some analytical explanations, it is of some interest to compare numerical results obtained from these approximations, with 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, the results from predictions and iterations are strongly correlated. Despite the differences, these results appear to be sufficiently close to justify the application of the modified inverse elasticity rule for analytical purposes. In the following discussion, when working out extensions for various factors, we use the modified inverse elasticity rule without income tax (A.11), because it is simpler. The same expressions would be obtained by applying the rule with income tax (A.15).

It should be noted that the accuracy of predictions also depend on the level of fixed public expenditure requirements. The predictions in table A.1 are based on public expenditure of 10%. It appears that in the present model, when public expenditure \( G \) differs from 0.1, using an adjustment factor defined as: \((G-0.1)^*0.7 +1\), for multiplying all predicted tax rates, will lead to much better approximations.
### Table A.1
COMPARING TAX RATES FROM ITERATIONS WITH PREDICTIONS FROM THE MODIFIED INVERSE ELASTICITY RULE

<table>
<thead>
<tr>
<th>Good</th>
<th>Without income tax</th>
<th>Without income tax</th>
<th>With income tax</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inequality aversion = 1</td>
<td>Inequality aversion = 0.3</td>
<td>Inequality aversion = 1</td>
</tr>
<tr>
<td></td>
<td>iterations predictions</td>
<td>iterations predictions</td>
<td>iterations predictions</td>
</tr>
<tr>
<td>1</td>
<td>0.98 1.15</td>
<td>0.32 0.38</td>
<td>0.42 0.80</td>
</tr>
<tr>
<td>2</td>
<td>0.85 0.67</td>
<td>0.30 0.24</td>
<td>0.34 0.24</td>
</tr>
<tr>
<td>3</td>
<td>0.22 0.35</td>
<td>0.11 0.18</td>
<td>-0.10 0.13</td>
</tr>
<tr>
<td>4</td>
<td>0.99 0.63</td>
<td>0.32 0.21</td>
<td>0.42 0.14</td>
</tr>
<tr>
<td>5</td>
<td>0.48 0.31</td>
<td>0.22 0.13</td>
<td>0.08 -0.04</td>
</tr>
<tr>
<td>6</td>
<td>0.98 0.66</td>
<td>0.32 0.22</td>
<td>0.42 0.18</td>
</tr>
<tr>
<td>7</td>
<td>0.45 0.42</td>
<td>0.21 0.18</td>
<td>0.07 0.08</td>
</tr>
<tr>
<td>8</td>
<td>0.68 0.53</td>
<td>0.27 0.21</td>
<td>0.22 0.13</td>
</tr>
<tr>
<td>9</td>
<td>0.43 0.45</td>
<td>0.21 0.20</td>
<td>0.06 0.11</td>
</tr>
<tr>
<td>10</td>
<td>1.49 2.02</td>
<td>0.39 0.30</td>
<td>0.74 0.67</td>
</tr>
<tr>
<td>11</td>
<td>2.68 2.80</td>
<td>0.46 0.36</td>
<td>1.50 1.23</td>
</tr>
<tr>
<td>12</td>
<td>1.62 2.05</td>
<td>0.40 0.30</td>
<td>0.83 0.71</td>
</tr>
<tr>
<td>13</td>
<td>1.55 2.03</td>
<td>0.40 0.30</td>
<td>0.78 0.69</td>
</tr>
<tr>
<td>14</td>
<td>1.32 1.96</td>
<td>0.38 0.28</td>
<td>0.63 0.60</td>
</tr>
<tr>
<td>15</td>
<td>2.12 2.41</td>
<td>0.43 0.35</td>
<td>1.15 0.99</td>
</tr>
<tr>
<td>16</td>
<td>2.34 2.51</td>
<td>0.45 0.35</td>
<td>1.29 1.06</td>
</tr>
<tr>
<td>17</td>
<td>1.31 1.95</td>
<td>0.38 0.28</td>
<td>0.62 0.59</td>
</tr>
<tr>
<td>18</td>
<td>1.70 2.12</td>
<td>0.41 0.31</td>
<td>0.88 0.77</td>
</tr>
<tr>
<td>Average</td>
<td>0.94 0.98</td>
<td>0.31 0.27</td>
<td>0.39 0.39</td>
</tr>
</tbody>
</table>

### Compliance costs

We start the discussion with compliance costs that are borne by the taxpayer. The definition of these costs in section 3.5.1 eq. (19) is: $c_i = c_ip_iq_i\theta_i$. Taking the indirect utility definition of the social welfare function in (1), then the dead-weight costs defined in (19) can be incorporated into the social welfare function by subtracting them from lump-sum income ‘y’. The extended social welfare function will be:

$$U = \sum_h a_h u_h(P, W_h, [y_{oh} - \sum_i c_i q_{ih}\theta_i])$$  \hspace{1cm} (A.16)

Note, $p_i$ has been omitted in (A.16), because all producer prices have been set to one. Applying the modified inverse elasticity rule to (A.16), we may notice that the presence of $c_i$ has a first order effect only on the personal utility term shown in (A.1). Deriving (A.16) with respect to $t_i$ we obtain:

$$\sum_h \frac{\partial u_h}{\partial t_i} = -\sum_h q_{ih} \frac{\partial u_h}{\partial y} - \sum_h \frac{\partial u_h}{\partial y} \left( c_i q_{ih} + t_i c_i \frac{\partial q_{ih}}{\partial t_i} \right)$$  \hspace{1cm} (A.17)
For the sake of brevity, we omitted $a_i$ from (A.17), and the same applies in the following equations. Combining (A.17) with the unchanged (A.6) and (A.8) terms and performing the operations from (A.4) to (A.10), we arrive at the following extended version of the modified inverse elasticity rule:

$$\begin{align*}
t_i &\approx 1 + t_i + \left[(1 + t_i) + \bar{e}_i t_i\right]c_i - \bar{u}_{mi}\left(1 + t_i + \bar{e}_i(t_i - t_{iA}) + \bar{e}_{it} \bar{T}_i^t\right) \overline{\bar{e}_i(u_0)} \\
\text{(A.18)}
\end{align*}$$

Subtracting (A.11) from (A.18) yields the estimated impact of $c_i$ on the optimal tax rate:

$$\begin{align*}
\Delta t_i(c_i) &\approx \frac{1 + t_i(1 + \bar{e}_i)c_i}{\bar{e}_i(u_0)} \\
\text{(A.19)}
\end{align*}$$

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:

$$R = \sum_h \int_{l} (I_{ih} - q_{ih} t_i s_i) - R_0$$

Deriving (A.20) with respect to $t_i$, we obtain a revised transfer term instead of (A.6)

$$\begin{align*}
\frac{\partial R}{\partial b} \frac{\partial R}{\partial t_i} &= \frac{\partial R}{\partial b} \left[\sum_h \left(q_{ih} - s_i \left(q_{ih} + t_i \frac{\partial q_{ih}}{\partial p_{il}} + \frac{\partial q_{ih}}{\partial p_{il}} \left(t_i - t_{iA}\right) + \frac{\varepsilon_{ih}(l)m}{p_{il}} T_{ih}'\right)\right]\right] \\
\text{(A.21)}
\end{align*}$$

Combining (A.21) with (A.1) and (A.8) and performing the operations from (A.6) to (A.10), we arrive at the following extended version of the modified inverse elasticity rule:

$$\begin{align*}
t_i &\approx \frac{1 + t_i - \bar{u}_{mi}\left(1 + t_i - s_i \left[(1 + t_i) + \bar{e}_i t_i\right] + \bar{e}_i(t_i - t_{iA}) + \bar{e}_{it} \bar{T}_i^t\right)}{\bar{e}_i(u_0)} \\
\text{(A.22)}
\end{align*}$$

Subtracting (A.11) from (A.22) yields the estimated impact of $s_i$ on the optimal tax rate:

$$\begin{align*}
\Delta t_i(s_i) &\approx \frac{s_i \bar{u}_{mi}\left[(1 + t_i) + \bar{e}_i t_i\right]}{\bar{e}_i(u_0)} \\
\text{(A.23)}
\end{align*}$$

**Externalities**

The starting point for deriving the formula for externalities is the extended social welfare function defined by (32) in section 3.5.4. In order not to complicate the analysis with cross-price effects, we shall concentrate here only on a single externality. In this case, the extended social welfare function will be:

$$U = \sum_h a_h u_h(P, W_h, \{y_{oh} + z_h \mu_i Q_i\})$$

where $Q_i = \sum_h q_{ih}$

Applying the modified inverse elasticity rule to (A.24), by deriving it with respect to $t_i$, we obtain:

$$\begin{align*}
\sum_h \frac{\partial u_h}{\partial t_i} &= - \sum_h q_{ih} \frac{\partial u_h}{\partial y} + \mu_i \sum_h \bar{z}_h \frac{\partial u_h}{\partial y} \frac{\partial Q_i}{\partial t_i} \\
\text{(A.25)}
\end{align*}$$

Notice that the externality term multiplying $\mu_i$ is opposite in sign to the utility derivative based on Roy’s Lemma. Combining (A.25) with the unchanged (A.6) and (A.8) terms and performing the operations from (A.6) to (A.10), we arrive at the following extended version of the modified inverse elasticity rule:


\[
1 + t_i - \bar{\mu}_i \frac{\Sigma_h z_h \frac{\partial u_h}{\partial y}}{\partial \bar{u}_i / \partial y} - \bar{u}_{mi}(1 + t_i + \bar{\epsilon}_i(t_i - t_{iA}) + \bar{\epsilon}_{iL} T_i')
\]

\[
t_i \approx \frac{- \bar{\mu}_i \Sigma_h z_h \frac{\partial u_h}{\partial y}}{\bar{\epsilon}_i(u_0) \frac{\partial \bar{u}_i}{\partial y}} (A.26)
\]

Subtracting (A.11) from (A.26), yields the estimated impact of \( \mu_i \) on the optimal tax rate:

\[
\Delta t_i(\mu_i) \approx - \frac{- \bar{\mu}_i \Sigma_h z_h \frac{\partial u_h}{\partial y}}{\bar{\epsilon}_i(u_0) \frac{\partial \bar{u}_i}{\partial y}} (A.27)
\]

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.13), multiplied by \( \bar{\epsilon}_i(u_0) \).

**Appendix 3.3**

**Further details about the setup of the numerical model**

In this appendix I present more details about the model described in section 3.3.

**C.1 Segmented utility**

First, we examine the segmented utility framework. As noted in section 3.3, in order to establish non-linear Engel curves using LES, nine commodities represent necessities and another nine with different utility parameters represent luxuries. LES being a linear Engel curves demand system, does not ensure automatically that goods defined by (4) will be non-negative. Non-negative consumption at low income levels (say when \( m = b \)) is possible only if the sum total of the positive intercept terms of goods (\( \alpha_i \) in eq. (4)) is smaller than \( b \), because the quantities demanded are always greater than \( \alpha_i \). Large negative intercepts are also not compatible with non-negative demand according to (4), when \( w \) and \( y \) are small. Thus, in order to ensure non-negative demand for all goods for consumers at low income levels, the intercept terms of necessities have to be small compared with average income. This has been confirmed by numerical results reported in Revesz (1997). But small intercept terms imply nearly homothetic preferences (when all \( \alpha_i = 0 \)), hence the utility function of necessities has to be nearly homothetic. With luxuries, the intercepts can assume larger positive or negative values without violating the non-negativity requirement, because the earning parameters of consumers (confined only to higher wage earners) are larger. Consequently, the utility function of higher wage earners can represent a wide range of income elasticities. Notice that in this model the consumption of luxury goods starts at intermediate income levels, which is not entirely unrealistic. This arrangement also ensures that nearly homothetic preferences are confined only to low wage earners.

Segmented LES represents non-identical preferences, but not heterogeneous preferences, as defined in Saez (2002a). Saez examines a model with multiple characteristics where consumers at the same income level choose different bundles of goods. In the present 15 person model, there is only one consumer at each income level. Nonetheless, segmented utility
does not accord with the assumption of identical preferences, adopted in the uniform commodity taxation theorems.

Revesz (1997) reported results from a 9 commodity LES model, where no segmentation was used, and all 15 taxpayers consumed only the 9 necessities and shared 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, but in order to avoid negative demand for any good, the 9 commodity linear Engel curves preferences had to be made nearly homothetic, which is not supported by empirical evidence. In this paper we are concerned only with results from the 18 commodity non-linear Engel curves model.

C.2 The cardinalisation of segmented utility

The application of two different utility functions for low and high wage earners raises some difficulties with cardinalisation. Since utility is an endogenous variable, the cardinal utility of the ninth taxpayer can sometimes turn out to be smaller than that of the eighth taxpayer, who has a lower wage rate. In other cases, there can be a large positive difference between the two. In order to ensure that utility is a smooth function of the wage rate, the program calculates and adds an adjusting constant to the utilities of all members of the higher wage group. Since the same constant is added to the utilities of all members of the second group in the uniform and non-uniform calculations, the utility outcomes in the two situations are comparable. However, the cardinalisation problem makes it difficult to carry out utility comparisons between different scenarios.

C.3 Linear income tax

Given zero homogeneous utility and demand in income and prices, in a model without income tax, a portion of commodity taxes can always be converted into a proportional income tax, without affecting utility or demand. Hence a model without income tax 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:

$$u(p, w, y) = u(1 + t_0, W, b)$$  \hspace{1cm} (C.1)

where \(t_0\) 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 (C.1), the demogrant ‘\(b\)’ also has to be reduced by ‘\(r\)’. After these adjustments \(u = u[(1 + t_0)(1 - r), W(1 - r), b(1 - r)]\), which yields exactly the same outcomes as \(u\) in (C.1), because of zero homogeneous utility and demand in prices and earning parameters. Following these changes, the commodity tax rates will be: \(t_i = (1 + t_0)(1 - r) - 1\) and the linear income tax function will be: \(T = rW - rb\). 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 (with \(b\) included) is actually a commodity cum linear income tax model and not purely a commodity tax model.
C.4 Inequality aversion rate

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 could be examined by letting the $a_h$’s to represent different political value judgments. Here we shall follow an inequality aversion rate 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, $g_i = c/u_{mi0}$. 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:

$$U = \sum_i u_i z + u_i (1 - z) c/u_{mi0}$$  \hspace{1cm} (C.2)

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 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) has been also examined.

C.5 Compensated elasticity of labour supply

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, but the dispersion between commodity tax rates is positively related to it. This rule is generally accepted in the literature. In regard to commodity taxation refer to the findings reported in section 2.3.2. In regard to income taxation, refer to Stern (1976), Revesz (1989) and eq. (1.5) in appendix 2.1. While the compensated elasticity of labour supply is part of the model (see eq. (5)), little attention has been given here to study systematically the effect of this endogenous variable on optimal tax rates. The few numerical results obtained, when setting $\beta_l$ to different values, suggest that the lower is the average compensated elasticity of labour supply, the higher will be optimal indirect tax rates, and the smaller will be the dispersion between tax rates.
C.6 Parameters employed in the model

### Table C.1
**LIST OF UTILITY PARAMETERS AND WAGE RATES**

<table>
<thead>
<tr>
<th>good</th>
<th>Low income group $\beta_i$</th>
<th>$\alpha_i$</th>
<th>High income group $\beta_i$</th>
<th>$\alpha_i$</th>
<th>taxpayer</th>
<th>wage rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.065</td>
<td>58</td>
<td>0.040</td>
<td>81</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>0.078</td>
<td>4</td>
<td>0.037</td>
<td>32</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>0.037</td>
<td>28</td>
<td>0.015</td>
<td>-9</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>0.046</td>
<td>-2</td>
<td>0.030</td>
<td>13</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>0.087</td>
<td>-3</td>
<td>0.042</td>
<td>-29</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>0.044</td>
<td>1</td>
<td>0.029</td>
<td>15</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>0.047</td>
<td>1</td>
<td>0.014</td>
<td>4</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>0.057</td>
<td>2</td>
<td>0.025</td>
<td>13</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>0.099</td>
<td>1</td>
<td>0.026</td>
<td>13</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td>0.035</td>
<td>-11</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td>0.030</td>
<td>4</td>
<td>11</td>
<td>15</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td>0.022</td>
<td>-5</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td></td>
<td>0.026</td>
<td>-7</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td>0.035</td>
<td>-16</td>
<td>14</td>
<td>30</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td>0.055</td>
<td>0</td>
<td>15</td>
<td>42</td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td>0.033</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td>0.032</td>
<td>-15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td>0.034</td>
<td>-6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>leisure</td>
<td>0.44</td>
<td>-70</td>
<td>0.44</td>
<td>-70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table C.2
**LIST OF UTILITY PARAMETERS IN THE TWO COMPOSITE GOODS MODEL**

<table>
<thead>
<tr>
<th>low income group</th>
<th>high income group</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_i$</td>
<td>$\alpha_i$</td>
</tr>
<tr>
<td>all goods 1 - 9</td>
<td>0.0622</td>
</tr>
<tr>
<td>all goods 10 - 18</td>
<td></td>
</tr>
<tr>
<td>leisure</td>
<td>0.44</td>
</tr>
</tbody>
</table>
Appendix 3.4
Supplementary tables

Table D.1
OPTIMAL COMMODITY TAX RATES WITH PROGRESSIVE INCOME TAX
Extension of Table 2

<table>
<thead>
<tr>
<th>good</th>
<th>Inequality aversion = 1</th>
<th>Inequality aversion = 0.3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>marginal utility ratios</td>
<td>compensated demand elasticity</td>
</tr>
<tr>
<td>1</td>
<td>1.13</td>
<td>-0.39</td>
</tr>
<tr>
<td>2</td>
<td>1.23</td>
<td>-0.71</td>
</tr>
<tr>
<td>3</td>
<td>0.91</td>
<td>-0.75</td>
</tr>
<tr>
<td>4</td>
<td>1.34</td>
<td>-0.84</td>
</tr>
<tr>
<td>5</td>
<td>1.08</td>
<td>-1.17</td>
</tr>
<tr>
<td>6</td>
<td>1.31</td>
<td>-0.79</td>
</tr>
<tr>
<td>7</td>
<td>1.05</td>
<td>-0.90</td>
</tr>
<tr>
<td>8</td>
<td>1.17</td>
<td>-0.82</td>
</tr>
<tr>
<td>9</td>
<td>1.04</td>
<td>-0.84</td>
</tr>
<tr>
<td>10</td>
<td>2.21</td>
<td>-1.27</td>
</tr>
<tr>
<td>11</td>
<td>2.08</td>
<td>-0.85</td>
</tr>
<tr>
<td>12</td>
<td>2.18</td>
<td>-1.20</td>
</tr>
<tr>
<td>13</td>
<td>2.19</td>
<td>-1.23</td>
</tr>
<tr>
<td>14</td>
<td>2.26</td>
<td>-1.43</td>
</tr>
<tr>
<td>15</td>
<td>2.11</td>
<td>-0.95</td>
</tr>
<tr>
<td>16</td>
<td>2.09</td>
<td>-0.91</td>
</tr>
<tr>
<td>17</td>
<td>2.26</td>
<td>-1.45</td>
</tr>
<tr>
<td>18</td>
<td>2.16</td>
<td>-1.13</td>
</tr>
</tbody>
</table>

Table D.2 shows average optimal tax rates corresponding to exogenously given demogrants (see section 3.4.4). The fixed demogrants were calculated using the average income level in the zero tax model. When taxation is present, due to the disincentive effect of the demogrant, average incomes will decrease, hence the demogrant to average income ratios will increase, as shown in the bottom row. The scenarios reported in Table D.2 are without income tax, expenditure on public goods amounts to 10% of total income under zero tax, and the inequality aversion rate is one.
Table D.2

OPTIMAL TAX RATES WITH SUB-OPTIMAL DEMOGRANTS

<table>
<thead>
<tr>
<th>Ratio of demogrant to average income in the pre-tax situation</th>
<th>0%</th>
<th>20%</th>
<th>40%</th>
<th>54%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average tax on necessities</td>
<td>0.05</td>
<td>0.22</td>
<td>0.46</td>
<td>0.70</td>
</tr>
<tr>
<td>Average tax on luxuries</td>
<td>0.22</td>
<td>0.78</td>
<td>1.56</td>
<td>1.83</td>
</tr>
<tr>
<td>Ratio of tax luxuries over necessities</td>
<td>4.4</td>
<td>3.5</td>
<td>3.4</td>
<td>2.6</td>
</tr>
<tr>
<td>% demogrant to average income</td>
<td>0</td>
<td>22</td>
<td>48</td>
<td>69</td>
</tr>
</tbody>
</table>

Table D.3 compares predicted optimal tax rates under tax evasion, with and without the adjustment factor discussed in section 3.5.3.

Table D.3

ACTUAL AND PREDICTED OPTIMAL TAX RATES WITH TAX EVASION

<table>
<thead>
<tr>
<th>With and without adjustment for dead-weight costs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Product No. and type</td>
</tr>
<tr>
<td>----------------------</td>
</tr>
<tr>
<td>Without income tax, inequality aversion = 1</td>
</tr>
<tr>
<td>Good 4 - necessity</td>
</tr>
<tr>
<td>Good 14 - luxury</td>
</tr>
<tr>
<td>Without income tax, inequality aversion =0.3</td>
</tr>
<tr>
<td>Good 4 - necessity</td>
</tr>
<tr>
<td>Good 14 - luxury</td>
</tr>
<tr>
<td>With income tax, inequality aversion = 1</td>
</tr>
<tr>
<td>Good 1 - necessity</td>
</tr>
<tr>
<td>Good 12 - luxury</td>
</tr>
</tbody>
</table>

Appendix 3.5

The mathematical framework with leisure complements and substitutes

This appendix examines two separate issues. First, we derive the modified inverse elasticity approximation for leisure complements and substitutes, presented in eq. (40) in the text. Following this, we outline the extended LES utility function that was used to obtain the numerical results reported in Tables 9 and 10.

In regard to the modified inverse elasticity rule, we may notice that among the three
components analysed in appendix 3.2, only the second component, that is tax revenue, is visibly affected by the presence of $\varepsilon_i$ 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 $\varepsilon_i$ from (38) as:

$$\varepsilon_i = \frac{\partial \bar{v}}{\partial p_i} = \frac{\partial \bar{v}}{\partial q_i} \frac{\partial q_i}{\partial p_i}$$  \hspace{1cm} (E.1)

Now, the revised revenue term (based on (A.6)) will be:

$$\frac{\partial u}{\partial b} \frac{\partial b}{\partial t_i} = \frac{\partial u}{\partial b} [\sum_h (q_{ih} + W_h T'_h \frac{\partial \bar{v}}{\partial p_i} + W_h T'_h \frac{\partial \bar{v}}{\partial q_i} \frac{\partial q_i}{\partial p_i} + \frac{\partial q_{ih}}{\partial p_i} (t_i - t_{i,A}))]$$  \hspace{1cm} (E.2)

$W_h$ is the wage rate of taxpayer h and $T'_h$ is the marginal indirect tax rate of consumer h (see A.5b). The first labour term in (E.2) refers to the labour supply derivative under weakly separable utility, as defined in (A.5a). The second term defines the effect of deviation from weak separability on the derivative of labour supply, as defined in (38). It represents the change in labour income due to leisure substitution or complementarity, multiplied by the marginal indirect tax rate, yielding the change in tax revenue due to the $\varepsilon_i$ factor.

Combining (E.2) with the unchanged (A.1) and (A.8) terms and performing the operations from (A.5) to (A.10), we arrive at the following extended version of the modified inverse elasticity rule:

$$t_i(\hat{\ell}) \approx \frac{1 + t_i - \bar{u}_{ml} \left(1 + t_i + \bar{c}_i \bar{m} + \frac{\partial \bar{v}}{\partial q_i} \bar{W}_i \hat{T}'_l + \bar{u}_i (t_i - t_{i,A})\right)}{\varepsilon_i(u_q)}$$  \hspace{1cm} (E.3)

Subtracting (A.11) from (E.3) and cancelling the similar price elasticity terms in the numerator and denominator, we obtain:

$$\Delta t_i(\hat{\ell}) \approx - \frac{\partial \bar{v}}{\partial q_i} \bar{u}_{ml} \hat{T}'_l \bar{W}_i$$  \hspace{1cm} (E.4)

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

**Extended LES**

To obtain the numerical results in Tables 9 and 10, we used an extended version of LES. With extended LES, labour supply in (5) is defined as:

$$\ell = Z - q_L = Z (1 - \beta_L) - \alpha_L - \frac{\beta_L}{w} (y - \sum_j p_j \alpha_j) - \sum_i k_i q_i$$  \hspace{1cm} (E.5)

where the $k_i$s 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:

$$u = \sum_i \beta_i \log (q_i - \alpha_i) + \beta_L \log (q_{L0} + \sum_i k_i q_i - \alpha_L)$$  \hspace{1cm} (E.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 (E.1) and (E.5), $\frac{\partial \bar{v}}{\partial q_i} = - k_i$

Obviously, the $k_i$s introduce an inter-dependency between commodities and leisure. It is not clear whether (E.6) can be solved to yield explicit global formulae for commodity demand and labour supply. Moreover, with the definition of utility in (E.6), it is unclear whether global quasi-concavity will be maintained. But even without global formulas, it is possible to derive local formulae of a given vector $q_{0i}$, obtained from the weakly separable utility model. Suppose the $k_i$s satisfy the following initial condition:
\[ \sum_i k_i q_{i0} = 0 \] \hspace{1cm} (E.7)

From (E.7) it follows that (E.5) will continue to satisfy zero homogeneous labour supply at \( q_{i0} \). 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 (E.5) his/her new choice of leisure will be:

\[ q_L = q_{L0} + \sum_i k_i (q_i - q_{i0}) \] \hspace{1cm} (E.8)

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

\[ W(\ell - \ell_0) = W\Delta \ell = \sum_i p_i (q_i - q_{i0}) = \sum_i p_i \Delta q_i \] \hspace{1cm} (E.9)

To determine \( \Delta q_i \), we employed the marginal consumption propensities, \( \partial q_i / \partial y \).

From the LES demand equation in (4):

\[ \partial q_i / \partial y = \beta_i / p_i \] \hspace{1cm} (E.10)

After taking out \( \beta_L \) and redefining the \( \beta_i \) as \( \tilde{\beta}_i = \beta_i / (1 - \beta_L) \), the sum will equal to one. Using this definition, we obtain from (E.10):

\[ \sum_i \frac{\partial q_i}{\partial y} p_i = \sum_i \frac{\tilde{\beta}_i}{p_i} p_i = \sum_i \tilde{\beta}_i = 1 \] \hspace{1cm} (E.11)

where \( \tilde{\beta}_i \) are the recalibrated values of \( \beta_i \) for commodities. To maintain proportionality between \( \partial q_i / \partial y \) and \( \Delta q_i \), and ensure that the budget constraint (E.9) is satisfied:

\[ \Delta q_i = W(\ell - \ell_0) \frac{\tilde{\beta}_i}{p_i} = W(\ell - \ell_0) \frac{\tilde{\beta}_i}{1 + \ell_i} \] \hspace{1cm} (E.12)

Adding the values found for \( \Delta \ell \) from (E.8) and for \( \Delta q_i \) from (E.12) to the original \( q_{L0} \) and \( q_{i0} \) variables, yields new values for commodities and leisure that can be used to evaluate the utility in (E.6). From then on, the calculations follow the procedure described in section 3.3.

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 \( q_{i0} \). 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 utility functions, such as those outlined in section 2.3, are better suited for testing the effect of leisure non-separability on optimal tax rates, but that challenge is left for future research.
References


Atkinson-Stiglitz theorem”, in Imperfect Economics: Essays in Honor of Joseph Stiglitz,


The Theory of Taxation and Public Economics”, Journal of Economic Literature, 48:4:
964–979.


Christiansen, V. (1984), “Which commodity taxes should supplement the income tax?”.


Cremer, H., Pestieau, P. and J. Rochet (2001), “Direct versus indirect taxation: the design of

Political Economy, 8: 299–312.


Deaton, A. (1979b): “The distance function in consumer behaviour with application to index

1260.


Deaton, A and J. Muellbauer (1980), Economics and consumer behavior, Cambridge:
Cambridge University Press.

Deaton, A. and N. Stern (1986), “Optimally uniform commodity taxes, taste differences and
lump-sum grants”, Economic Letters, 20: 263–266.

De Freitas, J. (2012). “Inequality, the politics of redistribution and the tax mix”, *Public Choice*, 151(3-4): 611-630.


OECD (2012). *Less Income Inequality and More Growth–Are They Compatible?*, OECD


Tuomala, M. (2016), *Optimal Redistributive Taxation*, Oxford University Press, United Kingdom


Post-script

Corrections to the discussion on the Atkinson-Stiglitz theorem

In section 2.1 I stated that the mathematical correctness of the Atkinson-Stiglitz (A-S, 1976) and Deaton (1979) theorems is not in doubt. Actually, there are some question marks about the mathematical correctness of the A-S theorem that has been discussed in Revesz (1986). The counterargument rests on the idea that even if the income tax function is the fully optimal Mirrleesian solution, further improvement can be made by adding differentiated commodity taxes to the optimal income tax solution.

To explain this idea a bit further, assume that the model includes egalitarian objectives, implying decreasing social marginal utilities of income. Then even when the variable part of all commodity tax functions is zero, as indicated by the A-S variational theorem under conditions of weak separability between goods and leisure, it is still possible to use lump-sum taxes for the right to purchase luxury goods (license fees), in order to raise social welfare higher. By the same token, social welfare could be improved by introducing differentiated demogrants, based on observable surrogate characteristics of ability (potential).

When commodity tax rates are proportional, and not in a general unspecified form, then Christiansen (1984) showed that optimal commodity tax rates should be all zero. Christiansen’s proof uses a Pollak leisure demand function, which is outside the realm of conventional demand theory. His argument is contradicted by the simpler derivation of optimality conditions on proportional tax rates in Revesz (1986), which shows that regardless of whether the income tax function is optimal or otherwise, further improvement is possible by introducing progressive commodity tax rates. More in-depth analysis of the proportional and variational solutions is presented in Revesz (1986).

The possible flaws in the A-S (1976) theorem may explain the inconsistency between the results in Revesz (2020) and the A-S theorem. Based on numerical results, Revesz (2020) finds that the optimal linear tax mix should be made up mainly of commodity taxation, whereas the A-S theorem states that the optimal non-linear tax mix should be made up entirely of labour income tax. Apart from the fact that one model deals with linear taxation and the other with non-linear taxation, the issue about the mathematical correctness of the A-S theorem could be behind the divergent results. It should be noted that the A-S (1976) theorem has never been tested with numerical results.

Even if one accepts the idea that the A-S theorem is entirely correct, there is still a major problem remaining. In section 2.2.2.1 I stated that the A-S (1976) theorem implies the optimality of zero or uniform commodity taxation. But in fact, the A-S (1976) theorem, and its generalisation by Christiansen (1984), only implies the optimality of zero commodity taxation in the presence of an optimal non-linear (Mirrleesian) income tax. It says nothing about the structure of optimal commodity tax rates in the presence of a sub-optimal income tax.
The theoretical support for uniform commodity taxation comes from a different source, namely the Deaton (1979) theorem. As explained in section 2.2.2.7, the Deaton theorem implies the optimality of uniform commodity taxation in the presence of linear Engel curves for all goods, a linear income tax and weak separability between goods and leisure. Yet, there is no empirical evidence to support the assumption of linear Engel curves for all goods. The fact that the uniform commodity taxation proposition is based on an even less realistic setting than the A-S theorem, has not been properly explained in the literature.

It is possible to combine uniform commodity taxation with a Mirrleesian income tax function (see Appendix 2.1). With a revenue neutral rearrangement, the outcome is likely to be better than with purely non-linear income tax optimisation. The reason is that with the introduction of uniform commodity taxation the demogrant will be also taxed. As a result, lump-sum income will be reduced, thus causing an increase in labour supply (refer to the discussion on eq. 15). However, such a model, involving negative marginal income tax rates near the end points, appears less politically acceptable than the purely income tax solution, which requires zero marginal income tax rates at the end points.

An additional reference

Buy your books fast and straightforward online - at one of world’s fastest growing online book stores! Environmentally sound due to Print-on-Demand technologies.

Buy your books online at
www.morebooks.shop

Kaufen Sie Ihre Bücher schnell und unkompliziert online – auf einer der am schnellsten wachsenden Buchhandelsplattformen weltweit! Dank Print-On-Demand umwelt- und ressourcenschonend produziert.

Bücher schneller online kaufen
www.morebooks.shop