The optimal grouping of commodities for indirect taxation[1]

Pascal Belan[2] Stéphane Gauthier[3] Guy Laroque[4]

January 25, 2008

[1]We are indebted to the editor Robin Boadway and to two anonymous referees for their constructive remarks. We are most thankful to Ian Crawford, from the Institute for Fiscal Studies, who computed the elasticities and budget shares in the U.K., which we use in the empirical application. We have benefited from the comments of Stuart Adam, Richard Blundell, Martin Browning, Jean-Marc Robin, Stephen Smith. We also thank the participants in seminars at CREST and IFS. The usual disclaimer applies.
[2]LEN, Université de Nantes.
[3]ENSAE and CREST.
[4]CREST-INSEE and University College London.

Abstract

Indirect taxes contribute to a sizeable part of government revenues around the world. Typically there are few different tax rates, and the goods are partitioned into classes associated with each rate. The present paper studies how to group the goods in these few classes. We take as given the number of tax rates and study the optimal aggregation (or classification) of commodities of the fiscal authority in a second best setup. The results are illustrated on data from the United Kingdom.

Les impôts indirects forment une part notable des recettes fiscales. D'ordinaire, on observe un petit nombre de taux différents, et les biens sont répartis en classes associées à chacun de ces taux. On étudie ici comment grouper les biens au mieux. Le nombre de taux est supposé fixé de manière exogène, et on résout le problème d'agrégation (ou de classement) optimal des biens dans un cadre de second rang. Les résultats sont illustrés sur des données britanniques.

Keywords: indirect tax, Ramsey, aggregation.
JEL classification numbers: H21, H23

1 Introduction

The French government recently wanted to change the rate of the value added tax bearing on meals taken in restaurants, but the European Union did not accept France's demand. The standard theory of indirect taxation would possibly recommend to tax restaurants at a higher rate than fast food places, e.g. because rich households spend a larger fraction of their income in restaurants than the less well-off. This theory, however, does not take into account a strong constraint imposed on EU members. Indeed, according to the 92/77 directive, EU members are allowed to set only one or two reduced (low) rates in addition to the standard (high) tax rate, so that they are forced to impose the same rate on many different commodities. The purpose of this paper is to describe how different commodities should be grouped when there is a constraint on the number of tax rates.

The Ramsey tax rule usually assumes that different commodities can be taxed at different rates, and when consumers have heterogeneous tastes and income, the optimal tax rates typically differ across goods.[1] In a partial equilibrium framework, assuming no substitution between goods, each commodity is assigned two numbers: its demand elasticity with respect to own price and its social weight, which reflects its relative usage among the poor and wealthy in the population. For a given social weight, the optimal tax rate is inversely proportional to the price elasticity; and, given the elasticity, the tax rate decreases with the social weight.

The situation where the number of available tax rates is smaller than the number of taxable goods has received little attention in the literature. Two previous theoretical papers are relevant. Gordon (1989) studies an economy where all goods are initially taxed at the same rate, and considers small changes in a tax reform perspective. Belan and Gauthier (2004) and Belan and Gauthier (2006) study the case of low (close to zero) levels of collected tax in a single agent framework with a finite number of goods. They find that the optimal tax rate bearing on a good is weakly decreasing in the price elasticity. To the best of our knowledge, there is no applied analysis of this issue.

In this paper, we depart from Belan and Gauthier (2006) by considering a continuum (instead of a finite number) of taxable commodities. Each good is

[1]See Saez (2002). When the consumers have the same tastes and when their labor supplies are separable from their demands for commodities, nonlinear income taxation yields a uniform taxation of all the goods (Atkinson and Stiglitz (1976) and Mirrlees (1976)). Apart from non separability and/or heterogeneous tastes, optimal indirect tax rates may differ across commodities for certain types of production functions (Stiglitz (1982), Naito (1999) or Saez (2004)), if it is possible to evade tax (Boadway, Marchand, and Pestieau (1994)), in order to correct externalities (Green and Sheshinski (1976)), in presence of uncertainties (Cremer and Gahvari (1995)), or when the authority implementing direct taxes is not perfectly coordinated with the one that designs indirect taxes, possibly because the decisions are taken at different points in time or in space (federal, state or city levels).

assumed to be negligible with respect to the total, so that it is possible to change the tax rate bearing on an elementary commodity while leaving unchanged the whole tax structure and the marginal cost of public funds. This allows us to consider arbitrary (far from zero) levels of taxes and to derive simple properties of the optimum. We also take into account heterogeneity and equity concerns. The theoretical predictions are used to study whether actual tax systems depart from optimality. We give a first look at this topic on data from the UK. The assumption that the observed tax rates are optimal provides information on the underlying social welfare function, which in turn puts restrictions on how to tax the goods.

The continuum assumption allows us to partially characterize the optimal grouping structure with the help of a *purported tax rate*. Such a rate is defined as the one that the social planner would apply to a good if this good could be taxed freely, while keeping unchanged the tax rates supported by the other commodities, fixed at their constrained optimum values. If social welfare satisfies a single-peakedness property with respect to tax rates, the highest attainable welfare must be at one of the typically two rates which are the closest to the ideal rate, on either side of this rate. It remains some uncertainty, at this stage, concerning the exact conditions under which a given good should be taxed at one particular rate, among the two relevant candidates.

In practice, it is unlikely that reliable information be available on how price elasticities change with prices. If one focuses attention on the particular case where the price elasticities and social weights do not vary with the tax rates, then one can fully characterize the tax structure, and answer the question about the exact conditions under which any individual good should be taxed at a given rate. The result is clear-cut: the Ramsey monotonicity properties are shown to be weakly satisfied. That is, given the price elasticity, the tax rate is non increasing with the social weight, and similarly, given the social weight, the tax rate is typically non increasing with the price elasticity.

The previous argument provides insights on whether a given actual fiscal scheme is optimal and why it may not be. In order to apply these results to data from the United Kingdom, we extend the analysis to the case in which cross price effects are not zero. We assume that the observed rates on the existing groups are optimally chosen. This yields constraints on the implicit redistributive aims of the government. It appears that the social weights that best fit the current tax scheme put most of the weight on the population segment associated with the middle of the consumption distribution, the fourth and fifth deciles. For these social weights, the actual commodity groupings do not look far from optimality. The main departures concern goods whose taxation is likely to rely on other considerations, environment or public health considerations, than mere redistribution. Thinking of the French demand to the European Union, in the U.K., 'Food out', which comprises restaurants and fast food places, is currently taxed at the standard rate, but appears to be too heavily taxed. Our analysis actually suggests that some items in this group should be exempted from any tax.

The paper is organized as follows. A simple framework which assumes separability between goods is first analyzed. It is described in Sections 2 and 3, and the optimal indirect tax schedule when the number of tax rates is restricted is characterized in Section 4. Section 5 extends the results to the case in which cross price effects matter. An application to data from the United Kingdom is presented in the final section. The proof of the main result is at the end of the paper. An Appendix with supplementary material appears on the web site of the journal.

2 Consumers

There is a continuum of goods g, g in G, and a numeraire. The typical consumer, designated with an index c, c in C, maximizes

$$\int_G u(x_g, g, c)\mu(g)dg + m$$

under her budget constraint

$$\int_G (1+t_g)x_g\mu(g)dg + m = w_c.$$

Both sets C and G are equal to the $[0,1]$ interval of the real line. The function u, defined over $\mathbb{R}_+ \times G \times C$, is assumed to be increasing, concave and twice continuously differentiable with respect to consumption x_g, x_g in $\mathbb{R}_+$, and continuous with respect to g and c. The consumption of numeraire is denoted by m. The relative importance of the various commodities is partially captured by their density μ with respect to the Lebesgue measure. The units of commodities are chosen so that all producer prices equal 1. Thus, when commodity g is taxed linearly at rate t_g, $t_g \geq -1$, the consumer price is $1 + t_g$. Finally, w_c is the exogenous income of consumer c.

Under the usual Inada conditions, the separability assumptions embodied in (2) imply that the demand $\xi_g(t_g, c)$ of commodity g by consumer c is the unique solution of the first-order condition $u'_x(x, g, c) = 1 + t_g$. It is decreasing and continuously differentiable with respect to the tax rate.

The indirect utility from consuming a good g taxed at rate t_g, writes

$$v_g(t_g, c) = u(\xi_g(t_g, c), g, c) - (1+t_g)\xi_g(t_g, c),$$

and therefore the overall indirect utility of consumer c is

$$\int_G v_g(t_g, c)\mu(g)dg + w_c.$$

3 Optimal unconstrained tax schedules

When choosing indirect taxes, the government takes as given market behavior. It seeks to maximize the sum of the utilities of the consumers, weighted by some *a priori* weights α_c, $\alpha_c \geq 0$ for all c, normalized so that

$$\int_C \alpha_c d\nu(c) = 1,$$

where ν is the (probability) measure describing the distribution of consumer characteristics on the set C.

The objective of the government can be written as the sum

$$\int_G V_g(t_g)\mu(g)dg, \tag{1}$$

where

$$V_g(t_g) = \int_C \alpha_c v_g(t_g, c)d\nu(c).$$

By Roy's identity,

$$\frac{dV_g}{dt_g}(t) = -a_g(t)X_g(t),$$

where

$$X_g(t_g) = \int_C \xi_g(t_g, c)d\nu(c)$$

represents the aggregate demand for good g, and

$$a_g(t) = \int_C \alpha_c \xi_g(t, c)/X_g(t)d\nu(c)$$

is a positive number which measures the *social weight of good g*. Namely, it is large when the agents c with the largest weights α_c consume relatively more of the good.

Assume first that there is no constraint on rates setting. When fiscal income to be collected is R, the government has to choose tax rates t_g, $g \in G$, which maximize (1) under the budget constraint

$$\int_G t_g X_g(t_g)\ \mu(g)dg = R.$$

Let λ denote the multiplier associated with the budget constraint. At the optimum, it corresponds to the marginal cost of public funds. The government problem is equivalent to maximizing

$$\int_G \mathcal{L}_g(t_g)\mu(g)dg,$$

where the Lagrangian $\mathcal{L}_g(t_g) = V_g(t_g) + \lambda t_g X_g(t_g)$ represents the contribution of good g to the welfare objective. If the authority can freely choose the tax rate bearing on good g, the necessary first-order condition for an interior optimum is, appealing to Roy's identity, $-a_g(t)X_g(t) + \lambda\left(X_g(t) + tX_g'(t)\right) = 0$ or, dropping the index g to simplify notations,

$$\frac{t}{1+t} = \frac{\lambda - a}{\lambda}\frac{X}{-(1+t)X'}. \tag{2}$$

This corresponds to the Ramsey rule, in which the tax rate applying to a consumption good is inversely related to the price elasticity $-(1+t)X'/X$ of the (aggregate) demand for this good[2].

One should be careful when using the first order condition since the program of the government is not well behaved, the Lagrangian $\mathcal{L}_g(t_g)$ being often not concave in t_g.

For instance, in the case where all the agents have demands for good g with the same constant price elasticity ε_g, a necessary and sufficient condition for global concavity of the Lagrangian is $\lambda(1-\varepsilon_g) \geq a_g$ (see Appendix A on the journal web site). If $\lambda(1-\varepsilon_g) < a_g$, then the Lagrangian is first concave and then convex. With constant price elasticity nevertheless, whatever the direction of the inequality, the Lagrangian is single peaked, so that the first order condition characterizes a global maximum.

In order to encompass such situations, we use the following assumption:

Assumption 1 *Given the marginal cost of public funds λ, a good g satisfies the* single peaked assumption *when the function $\mathcal{L}_g$, defined on $(-1, +\infty)$, satisfies one of the following three properties:*

1. *It is increasing;*
2. *It is increasing from -1 to some $\tau_g(\lambda)$ and decreasing from then on;*
3. *It is decreasing.*

There is another large class of situations, on top of the constant elasticity case, where the single peaked assumption is easy to check. This is when the elasticity of aggregate demand with respect to the tax rate, tX'/X, is non increasing in t, and the social weight of the good is non decreasing[3]. Contrary to concavity, however,

[2]Indirect taxation is useless when $\lambda = a_g$ for all consumption goods g, a condition unlikely to be satisfied when the agents do not have the same tastes, as emphasized in the recent literature, e.g. in Saez (2002).

[3]The social weight of a good is non decreasing in the tax rate when the demand of the socially unfavored (rich) agents decreases relatively to that of the socially favored (poor) agents when the tax increases. To show the property, observe that the derivative $\mathcal{L}' = -aX + \lambda(X + tX')$ has the same sign as $(\lambda - a)/\lambda + tX'/X$. Thus, if a is non decreasing in t and tX'/X is non increasing, $\mathcal{L}'$ at most has one change of sign, so that $\mathcal{L}$ is single peaked.

the single peakedness property is not preserved under aggregation: the sum of single peaked functions is not always single peaked (see Figure 3 of Appendix A).

The normal situation is that of Assumption 1.2. The analysis is easily extended when the solution goes to the boundaries of the tax domain: under Assumption 1.1 (resp., Assumption 1.3), the optimal tax rate is equal to $+\infty$: the good is made infinitely expensive (resp., to -1: the good is made free).

4 Tax rule with a finite number of rates

Assume now that there is an *a priori* given finite number K of different tax rates, t_k, $k = 1, \dots, K$, without loss of generality ranked in increasing order, $t_k \leq t_{k+1}$ for all k. Let G_k be the subset of goods which are taxed at rate t_k and $\mathbf{G}$ the collection of G_k. The government program becomes:

$$\begin{cases} \max\limits_{(t_k, G_k)_{k=1}^K} \sum_{k=1}^{K} \int_{G_k} V_g(t_k)\mu(g)dg \\ \sum_{k=1}^{K} \int_{G_k} t_k X_g(t_k)\ \mu(g)dg = R \\ \bigcup_{k=1}^{K} G_k = G. \end{cases} \tag{3}$$

The government has to choose the K tax rates (or possibly $K-1$, if one of them is constrained to be equal to zero) and the partition of the set of commodities associated with the various tax rates. Formally, this is a more complicated problem than the Ramsey problem, since it involves the variables $\mathbf{G}$, to which the standard Lagrangian methods do not immediately apply.

4.1 Optimal tax rates for a given partition of the goods

Given the partition $\mathbf{G}$, however, the problem is standard. Under usual regularity conditions, one can write the Lagrangian associated to this problem. When differentiating with respect to the tax rates, it is natural to consider the *aggregate* commodity G_k,

$$X_{G_k}(t) = \int_{G_k} X_g(t)\mu(g)dg.$$

The necessary first-order condition corresponding to t_k, first derived in Diamond (1973), can then be written as

$$\frac{t_k}{1+t_k} = \frac{\lambda - a_{G_k}}{\lambda} \frac{X_{G_k}}{-(1+t_k)X'_{G_k}} = \frac{\lambda - a_{G_k}}{\lambda} \frac{1}{\varepsilon_{G_k}}. \tag{4}$$

In this expression, a_{G_k} is the average of the social weights of the individual commodities in G_k,

$$a_{G_k} = \int_{G_k} \frac{X_g(t_k)}{X_{G_k}(t_k)} a_g(t_k)\mu(g)dg,$$

and ε_{G_k} is a weighted sum of the elementary price elasticities of the goods g in G_k,

$$\varepsilon_{G_k} = \int \frac{X_g(t_k)}{X_{G_k}(t_k)} \varepsilon_g(t_k)\mu(g)dg.$$

Hence, given the partition $\mathbf{G}$, optimal tax rates obey the Ramsey rule: The optimal tax rate t_k decreases with both the price elasticity ε_{G_k} of demand for goods in G_k and the social weight a_{G_k} of this group.

4.2 Optimal partition of the goods

In order to know how to allocate goods across groups, consider an individual commodity g, small with respect to the economy. Under the continuum hypothesis, a change in its tax rate leaves the marginal cost of public funds λ unchanged, since this parameter depends on the whole tax structure.

In this circumstance, we have:

Theorem 1 *A necessary condition for optimality is that, for almost every good, good g be attached to group G_k such that*

$$\mathcal{L}_g(t_k) = \max_{h=1,\dots,K} \mathcal{L}_g(t_h). \tag{5}$$

The assumption that there is a continuum of commodities, each negligible with respect to the whole economy, is crucial here. If a commodity were not of negligible size, as in Belan and Gauthier (2006), a change in the tax rate it supports would affect the marginal cost of public funds and the theorem would not hold. The result of the Theorem is quite intuitive, but its proof (at the end of the paper) needs some care, involving the Lyapunov theorem.

Let t_g^R be the tax rate that this good would support in the hypothetical situation where it would be taxed individually, all remaining goods being taxed at the (constrained) optimum. For an interior solution, t_g^R satisfies the Ramsey rule (2). A direct consequence of the single peakedness of the Lagrangian is

Lemma 1 *Under Assumption 1, at the optimum,*

1. *If $\mathcal{L}_g$ is increasing, good g belongs to the more heavily taxed group K; if it is decreasing, it belongs to the less taxed group.*

2. *Otherwise, with t_g^R the tax rate that maximizes $\mathcal{L}_g$,*

(a) if t_g^R is larger than t_K, commodity g supports the maximal rate;

(b) if there exists k, $k < K$, such that $t_k \leq t_g^R \leq t_{k+1}$, then g is taxed either at rate t_k or at rate t_{k+1};

(c) if t_g^R is less than t_1, g is taxed at rate t_1.

This lemma helps to describe some features of the optimal groups of commodities. Indeed, when only efficiency matters (a_g is identically equal to one for all g), Lemma 1 and the monotonicity of the Ramsey formula (2) in elasticities directly imply:

Theorem 2 *At an optimum, in the absence of redistribution motive, if the Ramsey price elasticity of good g, $\varepsilon_g^R = \varepsilon_g(t_g^R)$, is smaller than ε_{G_K}, then good g is taxed at the maximal rate t_K. If ε_g^R is larger than ε_{G_1}, then good g is untaxed. Otherwise, g is taxed at one of the k or $k+1$ rates such that*

$$\varepsilon_{G_k} \geq \varepsilon_g^R \geq \varepsilon_{G_{k+1}}.$$

In the absence of a redistribution motive, a weak version of the inverse elasticity rule consequently applies to individual goods, in the sense that the tax rate which should be supported by each individual good is non increasing with respect to its price elasticity, when evaluated at the putative free optimum.

In the more general case where the government has a redistributive objective, the social weights of the commodities typically differ from one. In the plan $(\varepsilon, a/\lambda)$, when the representative point $(\varepsilon_g^R, a_g^R/\lambda)$ of good g (with $a_g^R = a_g(t_g^R)$) belongs to the cone delimited by the two half lines

$$\frac{a}{\lambda} = 1 - \frac{t_k}{1+t_k}\varepsilon \quad \text{and} \quad \frac{a}{\lambda} = 1 - \frac{t_{k+1}}{1+t_{k+1}}\varepsilon,$$

Lemma 1 and the first-order condition (2) imply that it should be taxed at one of the two rates t_k or t_{k+1}.

Figure 1 is drawn with three tax rates, $t_1 = 0.0$, $t_2 = 0.2$, $t_3 = 0.4$: the three dashed lines are the corresponding half-lines which delimitate the cones. By Theorem 2, all goods g such that $a_g/\lambda \geq 1$ should be exempted, while all those such that $a_g/\lambda \leq 1 - t_3\varepsilon_g/(1+t_3)$ should be taxed at the highest rate $t = 0.4$.

Theorem 2 is not sufficient to pin down the exact optimal tax rate that should be applied to goods whose representative points stand in the intermediate region. Further insights can be gained from the special case with constant price elasticities and constant social weights. Then, one can indeed provide a full characterization of the partition of the plan. In the case of Figure 1 there are actually two curves, whose exact analytic expression is given in Appendix A on the journal web site, such that all goods in between these curves should be taxed at rate t_2, those

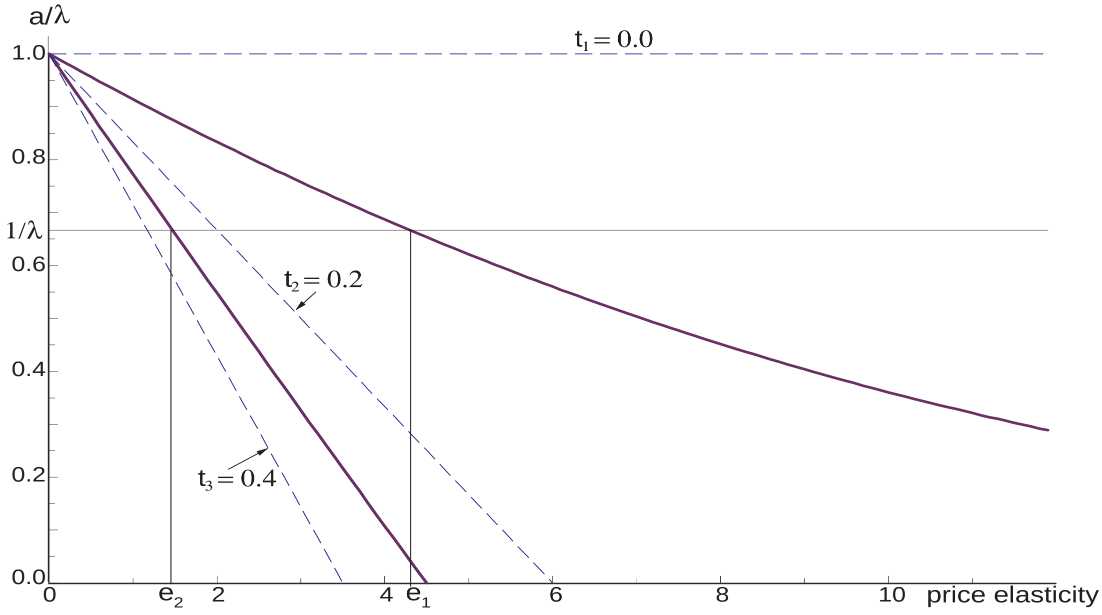

Figure 1: The efficient tax structure

above this region should be exempted, while those below should be taxed at the highest rate t_3. They are depicted in bold in Figure 1.

When tax rates are non-negative, theses curves are decreasing. In other words, in any given cone delimited by two half lines corresponding to two rates t and t', with $0 \leq t < t'$, there exists a unique threshold social weight above which a good should be taxed at the lowest rate, t, and below which it should be taxed at the highest rate, t'. This threshold appears to be decreasing with respect to price elasticity, which fits the common intuition: The larger the price elasticity of a good, the smaller the minimum social weight for which it becomes heavily taxed. Equivalently, for any social weight, there exists a threshold price elasticity such that a good in this cone will be taxed at the highest possible rate, t', if and only if its price elasticity is below this threshold.

Additional properties of these thresholds can be derived. In particular, the boundary curves are convex, and their slopes tend to that of the upper half line when the price elasticity goes to infinity. Strict convexity implies that, in any given cone, two commodities with the same purported ideal rates may be taxed at different rates in the optimal grouping.

To summarize, with constant price-elasticities and constant social weights, we have obtained a complete characterization of the tax structure at the level of individual goods. It turns out that a weak version of the Ramsey rule holds: The tax rate which should be applied to any individual good is weakly decreasing with respect to its own individual weight; when tax rates are non negative, it is also decreasing with respect to its own individual price elasticity.

5 Non separability of consumers' preferences

To apply our analysis to the data, we must enlarge the set of individual preferences and introduce, if possible, labor supply together with direct taxes. Let consequently the tastes of agent c be now represented by the utility function $U(\mathbf{x}, L_c, c)$, where $\mathbf{x}$ describes the consumption of goods, a measurable mapping from the set of commodities G into $\mathbb{R}_+$. The budget constraint of the typical consumer is:

$$\int_G (1+t_g)x_g\mu(g)dg = Y_c,$$

where Y_c is after tax income, i.e. $Y_c = w_cL_c - T(w_cL_c)$, for an income tax scheme T.

From now onwards, we shall work conditionally on the labor supply L_c. Let $\mathbf{t}$ be the collection of tax rates (t_g), and $V(\mathbf{t}, L_c, Y_c, c)$, the conditional indirect utility function of consumer c. The government chooses $\mathbf{t}$ which maximizes

$$\int_C \alpha_c V(\mathbf{t}, L_c, Y_c, c)d\nu(c)$$

subject to the budget constraint

$$\int_C \int_G t_g\xi_g(\mathbf{t}, L_c, Y_c, c)\mu(g)dg \quad d\nu(c) = R,$$

where $\xi_g(\mathbf{t}, L_c, Y_c, c)$ is the conditional (Marshallian) demand for good g of individual c.

Let ρ_c stand for the marginal utility of income of individual c. Using Roy's identity,

$$\frac{\partial V}{\partial t_g} = -\rho_c\xi_g(\mathbf{t}, L_c, Y_c, c),$$

and the necessary first-order condition for at maximum of the Lagrangian with respect to the tax rate t_k of group G_k is (see Appendix C on the journal web site)

$$\int_{g\in G_k}\left\{(-a_g+\lambda)X_g + \lambda\left[t_g\frac{\partial X_g}{\partial t_g} + \int_{g'\neq g} t_{g'}\frac{\partial X_{g'}}{\partial t_g}\right]\right\}\mu(g)dg = 0, \tag{6}$$

where the social weight of good g now expresses as

$$a_g = \int_C \frac{\xi_g}{X_g}\alpha_c\rho_c d\nu(c). \tag{7}$$

If good g could be taxed freely, the individual tax rate t_g^R would satisfy the first-order condition associated with an interior maximum,[4]

$$(-a_g+\lambda)X_g + \lambda\left[t_g^R\frac{\partial X_g}{\partial t_g} + \sum_{k=1}^{K} t_k\frac{\partial X_{G_k\setminus\{g\}}}{\partial t_g}\right] = 0. \tag{8}$$

[4]As indicated in Appendix C, equations (6) and (8) are obtained under the assumption that substitution between commodities is not too large.

The analysis of Section 4.2 can then be adapted to this more general setup. Namely, rewriting the first-order condition (8) as

$$\frac{a_g}{\lambda} - b_g = 1 - \frac{t}{1+t}\varepsilon_g, \tag{9}$$

where

$$b_g = \frac{1}{X_g}\sum_{k=1}^{K} t_k \frac{\partial X_{G_k \setminus \{g\}}}{\partial t_g}, \tag{10}$$

one can draw Figure 1 in the plan $(\varepsilon, a/\lambda - b)$ with a similar interpretation, provided that the Lagrangian is single peaked with respect to each tax rate t_g separately, with the representative point $(\varepsilon_g, a_g/\lambda - b_g)$ of good g evaluated at the optimal tax rates solution of (8).

6 Illustration with data from the UK

Professor Ian Crawford, from the Institute for Fiscal Studies, has provided us with uncompensated cross price elasticities for consumption in the UK, grouped into twenty categories[5], homogenous by tax rates, computed along the lines initiated by Blundell and Robin (1999), and with the budget shares by deciles of consumption expenditures in the population (the data is reproduced at the end of the Appendix). A large part of consumption, 49%, is subject to the 'standard' (17.5%) tax rate, and a substantial part, 27%, necessities including basic food, is either exempted or taxed at a zero rate. Our data do not separate exempted from zero rate items, and we treat the whole category as zero rated.[6] Domestic fuel, 10% of consumption, is taxed at the 'reduced' (5%) rate. Tobacco, alcohol, and petrol and diesel bear large excise tax rates.

In order to see whether the actual grouping of commodities fits the theory developed above, we assume that the tax authority takes as given after tax incomes and chooses optimally both the partition of the commodities and the tax rates. We want to check whether the data is consistent with this assumption. If VAT rates are optimally chosen, then they must satisfy the Diamond first order conditions (6). These restrictions provide some information on the underlying social weights used by UK government. Given these weights, one can compute the individual purported rates that would apply to each of the twenty categories of goods. This allows us to draw the analogous of Figure 1 for the UK, and thus to assess the optimality of the composition of the commodity groups.

[5] In the analysis, we shall drop 'children clothing', which represents less than 1% of aggregate consumption expenditure, because the estimated price elasticities are somewhat out of the ball park.

[6] The difference between the two categories comes from the fact that a producer of a zero rated good can reclaim the VAT bearing on his inputs, while the producer of an exempted good cannot. As a consequence, the exempted goods actually support some tax.

It is important to emphasize that such an exercise does provide information on the optimal indirect tax rates *given the current income tax schedule*. Indeed, if the government could freely tax income in a non linear way, and if, in addition, the Atkinson-Stiglitz conditions would prevail (preferences are separable between commodities and labor, and the preferences for commodities are identical across individuals at the microeconomic level), then all the goods should be taxed at the same rate (see Atkinson and Stiglitz (1976), Kaplow (2006), or Laroque (2005)). Here, however, we work with fixed after tax incomes.[7]

6.1 The government redistributive objectives

The redistributive stance of the government is represented by a vector of non negative weights associated with the ten population deciles, whose coordinates sum up to 1. For consistency, these weights should be such that the Diamond first-order conditions (6) for the basic three commodity groups, exempted, reduced rate and standard rate, are satisfied. The tax rates on alcohol, tobacco, and petrol do not give direct information on the redistributive stance, since they are likely to depend on other considerations than mere redistribution, e.g. public health or environmental issues.

As a result, given the observed tax rates, budget shares and price elasticities,[8] the ten unknown $\alpha_c \rho_c$ and the marginal social cost of public funds λ must satisfy four linear equations, the three first order conditions and the normalization condition. Therefore, in general, one cannot expect to recover the government objective from the Diamond first-order conditions: If there is some interior (strictly positive) solution to the equations, the set of solutions is locally a manifold of dimension $11 - 4 = 7$.

In practice, the minimum of the squares of the three left-hand sides of the Diamond conditions does nevertheless differ from zero. This gives a unique set of values for the ratios $(\alpha_c \rho_c / \lambda)$. If one normalizes the sum of $\alpha_c \rho_c$ over the deciles

[7] In addition, to the best of our knowledge, there is no general agreement on the empirical relevance of the Atkinson-Stiglitz conditions. Browning and Meghir (1991) find some evidence of non separability.

[8] Since there is a finite number of commodities, we have to rewrite (6) as

$$\sum_{g \in G_k} \left((-a_g + \lambda) X_g + \lambda \sum_{g' \in G} t_{g'} \frac{\partial X_{g'}}{\partial t_g} \right) = 0. \tag{11}$$

The consumption, rather than production, price is the numeraire. Using tildas for the variables measured with the new numeraire, $\tilde{X}_g = (1 + t_g) X_g$. After some manipulations, (6) becomes

$$(-a_{G_k} + \lambda) \frac{t_k}{1 + t_k} \tilde{X}_{G_k} + \lambda \sum_{k'} \frac{1}{1 + t_{k'}} \tilde{X}_{G_{k'}} \tilde{\varepsilon}_{G_{k'} G_k} = 0.$$

Finally, we work in shares of total consumption, dividing the equalities by total consumption.

to unity, and compute λ accordingly[9], one obtains

$$\hat{\lambda} = 1.11,$$

and most of the weight is on the fourth and fifth deciles

$$\alpha_1\rho_1 = 0.03 \qquad \alpha_4\rho_4 = 0.54 \qquad \alpha_5\rho_5 = 0.43.$$

The left-hand sides of the Diamond conditions (6) are respectively equal to 0.003 for the exempted goods, -0.007 for domestic fuel (the only good taxed at the reduced rate), and -0.0004 for goods taxed at the standard rate. These numbers are proportional, up to a positive factor, to the derivatives of the social objective with respect to the corresponding tax rates, such as given by (6). That is, they are equal to the social values of marginal changes of the tax rates, measured as tenths of aggregate consumption. For instance, increasing by 1 point the standard rate, from 17.5% to 18.5%, would induce a social loss of $0.0004 \times 0.01 \times 10 = 0.004\%$ of aggregate consumption. The Diamond first-order conditions are therefore close to be satisfied.

6.2 Is the grouping of commodities optimal?

Figure 2 plots the representative points $(\varepsilon_g, a_g/\lambda - b_g)$ of eighteen[10] commodities, and the half lines

$$\frac{a}{\lambda} - b = 1 - \frac{t}{1+t}\varepsilon,$$

corresponding to the current tax rates t. The parameters a, b and ε are set at their current observed values.[11]

In fact two points are drawn for each good. The one in large bold type corresponds to the implicit social weights computed above, while the other one, in small italic type, represents the good location for a Rawlsian government which would put all the social weight on the first population decile.

[9] This means that the social welfare function is normalized so that an increase of aggregate consumption of dC, uniformly distributed, gives $dC/10$ to each decile and therefore, for this choice of normalization, increases social welfare by $dC/10$. Thus, social welfare is implicitly measured in tenths of aggregate consumption.

[10] In the interest of readability, 'Wine and spirits' do not appear on the graph: its own price elasticity is (-)3, much larger than that of the other goods.

[11] Theory would require to compute the optimal putative tax rate for each considered commodity. For lack of better information, we assume that the observed elasticities are good enough approximations to be used to compute the graph coordinates. We have done some experimentation with more sophisticated computations for b and ε. In particular we have looked at cases where all the elasticities are constant, equal to their observed values, where demand functions are linear, and at a couple of other variants, including QAIDS which underlies the empirical estimation. The results are quite sensitive to the specification of the shape of the demand functions: in particular single peakedness is easily lost, and the Lagrangian may be locally convex at the observed point. More work is needed in this area.

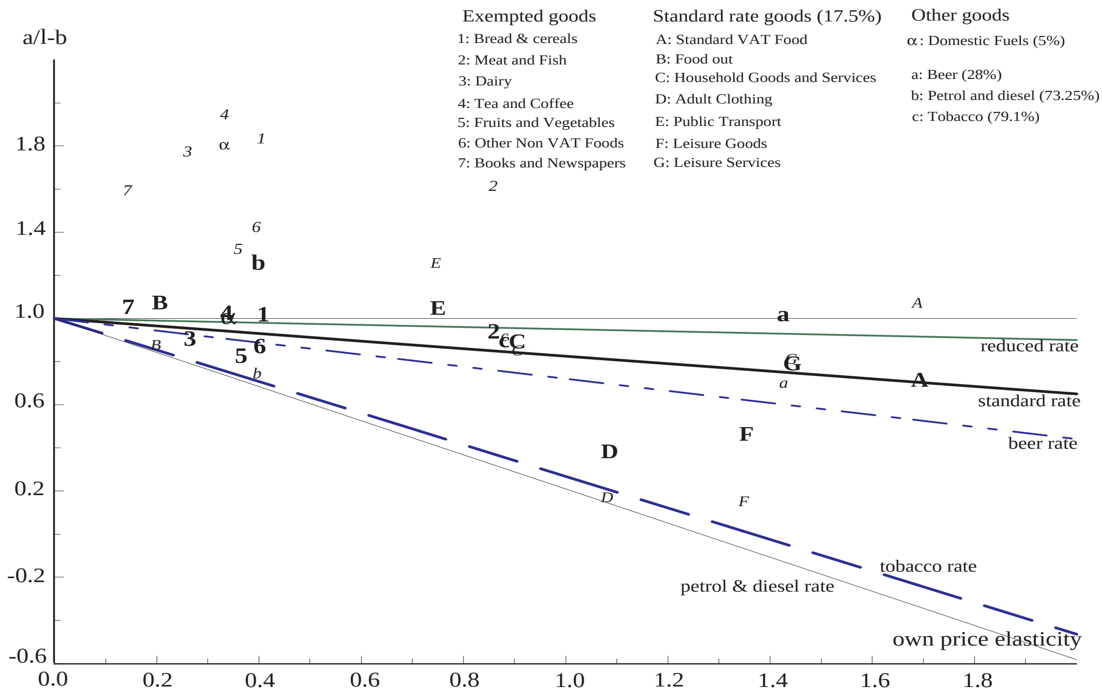

Figure 2: The fan for the UK

Under single peakedness, optimality requires that the large bold representative points of all the exempted goods be above the reduced rate half line, the point associated with 'Domestic Fuels' (the only good supporting the reduced rate) be between the standard rate line and the horizontal, and all the goods bearing the standard rate be below the reduced rate half line.

Excluding the goods subject to excise taxes, 87% of total consumption expenditures are concerned. Of these, 67% appear to be taxed consistently with the optimality criterion.[12] The main departures from optimality are the following. A number of exempted goods should be taxed at the standard rate: 'Dairy products', 'Fruits and Vegetables', and 'Other non VAT foods'. 'Food out' and 'Public transport', currently taxed at the standard rate, should be exempted.[13] At least in the U.K., if not in France, restaurants appear to be too heavily taxed.

If the government wants to raise more money by creating a larger tax rate, 'Adult Clothing' and 'Leisure Goods' seem to be good candidates to enter its basis.

[12]It is difficult to provide a statistical assessment of this result, given the possible measurement errors on elasticities. If one chooses, as a simple benchmark, the case in which the representative point of each class of goods were drawn independently uniformly in the half plan, given the half lines associated with each tax rate, then only 43% of the 87% consumption expenditures taken into account would be taxed according to the theory. This can be seen as a weak validation of the theory.

[13]Both 'Food out' and 'Public transport' are complementary with labor supply.

Four specific categories appear to be taxed more heavily than the redistributive social objective would recommend: 'Domestic Fuels', 'Beer', 'Petrol and Diesel' and 'Tobacco'. This may be justified on public health or environmental protection grounds[14]. The differences are large: for instance beer would be either exempted or taxed at a lower rate, and 'Domestic fuels' would be strongly subsidized.[15]

Going to a Rawlsian government allows us to look at the impact of the redistributive stance of the government. This tends to spread out the figure. A quarter of consumption ('Petrol and Diesel', 'Food out', 'Adult Clothing' and 'Leisure Goods') are taxed more heavily, a third ('Household Goods and Services', 'Leisure Services' and 'Tobacco') are unaffected, and the remainder, approximately 45% of consumption, gets a reduced rate or, more often a subsidy.

This rather surprising outcome indicates that differences in the consumption structure of the various deciles are large enough to make the optimal indirect tax rates vary substantially with the redistributive objective. In particular, the fact that consumers of the first decile devote a low fraction of their income to 'Food out', relative to the fourth and fifth deciles, implies that a Rawlsian social planner would heavily tax both restaurants and fast foods.

All things considered, these results look plausible and may be worth independent confirmation and further refinement.

References

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

Belan, P., and S. Gauthier (2004): "Optimal commodity grouping in a partial equilibrium framework," *Economics Letters*, 83(1), 49–54.

——— (2006): "Optimal indirect taxation with a restricted number of tax rates," *Journal of Public Economics*, 90(6), 1201–1213.

Blundell, R., and J. Robin (1999): "Estimation in Large and Disaggregated Demand Systems: an Estimator for Conditionally Linear Systems," *Journal of Applied Econometrics*, 14, 209–232.

[14] As suggested by a referee, it is likely that the results of Sandmo (1975) and Kopczuk (2003) extend to the present framework: an externality of the atmospheric type, caused by the total consumption of the polluting good, would be corrected by a Pigovian tax bearing only on this good.

[15] 'Wine and spirits' is subject to a 55% rate, but it should support a 36% rate, according to our computation.

Boadway, R., M. Marchand, and P. Pestieau (1994): "Towards a theory of the direct-indirect tax mix," *Journal of Public Economics*, 55(1), 71–88.

Browning, M., and C. Meghir (1991): "The Effects of Male and Female Labor Supply on Commodity Demands," *Econometrica*, 59(4), 925–951.

Cremer, H., and F. Gahvari (1995): "Uncertainty and optimal taxation: In defense of commodity taxes," *Journal of Public Economics*, 56(2), 291–310.

Diamond, P. (1973): "Taxation and public production in a growth setting," in *Models of economic growth*, ed. by J. Mirrlees, and N. Stern, pp. 215–235. International Economic Association, McMillan, London.

Gordon, J. (1989): "Tax reform via commodity grouping," *Journal of Public Economics*, 39, 67–81.

Green, J., and E. Sheshinski (1976): "Direct vs indirect tax remedies for externalities," *Journal of Political Economy*, 84(4), 797–808.

Hildenbrand, W. (1974): *Core and Equilibria of a Large Economy*. Princeton University Press.

Jahn, J. (2004): *Vector Optimization: Theory, Applications, Extensions*. Springer.

Kaplow, L. (2006): "On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal," *Journal of Public Economics*, 90(6-7), 1235–1250.

Kopczuk, W. (2003): "A Note on Optimal Taxation in the Presence of Externalities," *Economics Letters*, 80(1), 81–86.

Laroque, G. (2005): "Indirect Taxation is Superfluous under Separability and Taste Homogeneity : a Simple Proof," *Economic Letters*, 87, 141–144.

Mirrlees, J. (1976): "Optimal tax theory: A synthesis," *Journal of Public Economics*, 6(4), 327–358.

Naito, H. (1999): "Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency," *Journal of Public Economics*, 71(2), 165–188.

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

Saez, E. (2004): "Direct or indirect tax instruments for redistribution: short-run versus long-run," *Journal of Public Economics*, 88(3-4), 503–518.

Sandmo, A. (1975): "Optimal Taxation in the Presence of Externalities," *Swedish Journal of Economics*, 77.

Stiglitz, J. E. (1982): "Self-selection and Pareto efficient taxation," *Journal of Public Economics*, 17(2), 213–240.

Proof of Theorem 1

There is no vector space structure on the variables G, and therefore no way to differentiate with respect to G. To put a differentiability structure on the set of variables we abstract from the economic context and do as if it were possible to tax parts of good g at the various available rates. Let $\pi_k(g)$ be the fraction of good g subject to rate t_k, where $\boldsymbol{\pi}= [\pi_k(g), k = 1, \ldots, K]$, is a vector of positive measurable functions, defined on G, of square integrable with respect to the measure $\mu(g)dg$. The program (3) then becomes

$$\begin{cases} \max_{\mathbf{t},\boldsymbol{\pi}} \sum_{k=1}^{K} \int \pi_k(g) V_g(t_k)\mu(g)dg \\ \sum_{k=1}^{K} \int \pi_k(g) t_k X_g(t_k)\ \mu(g)dg = R \\ \pi_k(.) \geq 0, \text{ for all } k = 1, \ldots, K, \text{ and } \quad \sum_{k=1}^{K} \pi_k(.) = 1. \end{cases} \tag{3'}$$

where the variables maximized upon are $(\mathbf{t}, \boldsymbol{\pi})$ in $\mathbb{R}^K \times L_2^K(G)$ instead of $(\mathbf{t}, \mathbf{G})$. The only solutions of economic relevance are such that the functions π take only two values, either 0 or 1. An adaptation of the Lagrangian approach can be used to derive necessary conditions satisfied by a solution to the program (Theorem 7.3 of Jahn (2004)). Both the function to be maximized and the government revenue are Fréchet differentiable with respect to the variables $(\mathbf{t}, \boldsymbol{\pi})$. Let λ, $\rho = (\rho_k)$, and σ, respectively in $\mathbb{R}$, $L_2^K(G)$ and $L_2(G)$, be the multipliers associated with the government budget constraint, the positivity constraints and the normalization constraints. ρ is nonnegative and the solution is a local extremum of

$$\sum_{k=1}^{K} \int [\pi_k(g)V_g(t_k) + \lambda \pi_k(g) t_k X_g(t_k) + \pi_k(g)\rho_k(g) - \pi_k(g)\sigma(g)]\mu(g)dg,$$

with

$$\int \pi_k(g)\rho_k(g)\mu(g)dg = 0 \quad \text{for all } k,$$

and

$$\int [1 - \sum_{k=1}^{K} \pi_k(g)]\sigma(g)\mu(g)dg = 0.$$

Taking the Fréchet derivative with respect to π_k yields, for μ almost all g,

$$V_g(t_k) + \lambda t_k X_g(t_k) + \rho_k(g) - \sigma(g) = 0,$$

with $\rho_k(g) \geq 0$, $\pi_k(g) \geq 0$, and $\rho_k(g)\pi_k(g) = 0$. It follows that a necessary condition for optimality is

$$\sigma(g) = \max_{k=1,\ldots,K} [V_g(t_k) + \lambda t_k X_g(t_k)],$$

and that $\pi_\ell(g)$ is equal to zero whenever

$$\sigma(g) > [V_g(t_\ell) + \lambda t_\ell X_g(t_\ell)].$$

There are typically several optima, and there is always an economically meaningful solution in the set of optima, i.e. one solution such that $\pi_k(g)$ is everywhere either equal to 0 or to 1. This relies on the assumption that the space of commodities has no atoms, and directly follows from the following lemma:

Lemma 2 *Let Γ be a subset of goods such that, for $k = 1, \ldots, n$, there are real μ integrable functions α_k and β_k defined on Γ, verifying $\sigma(g) = \alpha_k(g) + \beta_k(g)$. Consider measurable functions from Γ into $[0, 1]$ such that $\pi_k(g)$, $\pi_k(g) \geq 0$, $\sum_k \pi_k(g) = 1$.*

Assume that the measure μ has no atoms on Γ. Then there exists a partition $(\Gamma_k)_{k=1,\ldots,n}$ of Γ such that:

$$A = \int_\Gamma \sum_k \pi_k(g)\alpha_k(g)\mu(g)dg = \sum_k \int_{\Gamma_k} \alpha_k(g)\mu(g)dg,$$

and

$$B = \int_\Gamma \sum_k \pi_k(g)\beta_k(g)\mu(g)dg = \sum_k \int_{\Gamma_k} \beta_k(g)\mu(g)dg.$$

Proof: For every g, let $\bar{\alpha}(g) = \max_k \alpha_k(g)$ and $\underline{\alpha}(g) = \min_k \alpha_k(g)$. Note also $\bar{k}(g)$ the smallest k such that $\bar{\alpha}(g) = \alpha_k(g)$, and similarly $\underline{k}(g)$ for the minimum. Of course:

$$\bar{A} = \int_\Gamma \bar{\alpha}(g)\mu(g)dg \geq A \geq \underline{A} = \int_\Gamma \underline{\alpha}(g)\mu(g)dg.$$

The non negative integral $\int_\gamma (\bar{\alpha}(g) - \underline{\alpha}(g))\mu(g)dg$, where γ is a measurable subset of Γ, defines a nonnegative atomless measure on Γ. By Lyapunov (see e.g. Hildenbrand (1974), p.45), its range is the convex interval $[0, \bar{A} - \underline{A}]$. There is therefore a set γ such that

$$A - \underline{A} = \int_\gamma (\bar{\alpha}(g) - \underline{\alpha}(g))\mu(g)dg.$$

For all k, define

$$G_k = \{g \in G | (g \in \gamma \text{ and } k = \bar{k}(g)) \text{ or } (g \notin \gamma \text{ and } k = \underline{k}(g))\}.$$

By construction the G_k's form a partition of G, and

$$\sum_k \int_{G_k} \alpha_k(g)\mu(g)dg = \int_\gamma \bar{\alpha}(g)\mu(g)dg + \int_{G\backslash\gamma} \underline{\alpha}(g)\mu(g)dg = A.$$

The second equality of the lemma is an immediate consequence of the equality $\alpha(g) = \sigma(g) - \beta(g)$.

■

The result follows from applying the lemma successively to all the subsets of tax rates $\chi = (k_1, ..., k_n)$ for which there exists a non-negligible set of goods such that $\mathcal{L}_k$ is constant on χ, and strictly smaller than $\mathcal{L}_{k_1}$ for k not in χ. The construction yields the same value of welfare and the same government receipts.

This completes the proof of the Theorem. Program (3) has more restrictive constraints than Program (3') and we have exhibited an admissible allocation for (3) that maximizes (3'). It satisfies the necessary conditions for optimality:

$$\mathcal{L}_g(t_k) = \max_{h=1,...,K} \mathcal{L}_g(t_h).$$

Appendix on 'The optimal grouping of commodities for indirect taxation'

A The case of constant elasticities

It is possible to derive a more precise characterization when all the consumers' demands have the same constant price elasticities for each good. The utility functions which yield demand functions whose price elasticities are constant are of the form

$$u(x,g,c) = \begin{cases} [A_g(c)]^{1/\varepsilon_g(c)} \, \dfrac{x^{1-1/\varepsilon_g(c)}}{1-1/\varepsilon_g(c)} & \text{for } \varepsilon_g(c) > 0,\ \varepsilon_g(c) \neq 1 \\ A_g(c) \ln \dfrac{x}{A_g(c)} & \text{for } \varepsilon_g(c) = 1. \end{cases}$$

The associated indirect utility functions are

$$v_g(t,c) = \begin{cases} \dfrac{A_g(c)}{\varepsilon_g(c)-1}(1+t)^{1-\varepsilon_g(c)} & \text{for } \varepsilon_g(c) > 0,\ \varepsilon_g(c) \neq 1 \\ -A_g(c) - A_g(c)\ln(1+t) & \text{for } \varepsilon_g(c) = 1. \end{cases}$$

The function $v_g(t,c) + \lambda t \xi_g(t,c)$ is not always well behaved. The second derivative of this function with respect to the tax rate is equal to

$$-\frac{\varepsilon}{(1+t)^{\varepsilon+2}}[(\lambda - A - \lambda\varepsilon)t + 2\lambda - A].$$

This expression is negative for t close to -1. The term in square brackets is increasing in t when $\lambda - A - \lambda\varepsilon$ is positive, so that the function is concave in this situation. Otherwise for $\lambda - A - \lambda\varepsilon < 0$, which always holds for elasticities larger than 1, there exists a $\hat{t}$ such that the function is concave on the interval $(-1, \hat{t}]$ and convex on $[\hat{t}, +\infty)$. In all circumstances, it is unimodal or single peaked.

The aggregate Lagrangian is

$$\mathcal{L}_g(t) = \int_C (\alpha(c) v_g(t,c) + \lambda t \xi_g(t,c))\, d\nu(c).$$

Single peakedness typically is not preserved by summation. An example is depicted on Figure 3, where there are two consumers with different price elasticities, and the aggregate demand is: $A_1(1+t)^{-\varepsilon_1} + A_2(1+t)^{-\varepsilon_2}$, with $\lambda = 1.25$, $\varepsilon_1 = 0.25$, $A_1 = 1.8$, $\varepsilon_2 = 4$ and $A_2 = 1$. Then, the tax rate t_g^R which maximizes $\mathcal{L}_g$ is between t_2 and t_3, but the optimal rate is t_1.

A sufficient assumption for the aggregate Lagrangian to be also single peaked, an assumption which is maintained in the rest of this Section, is that *the price*

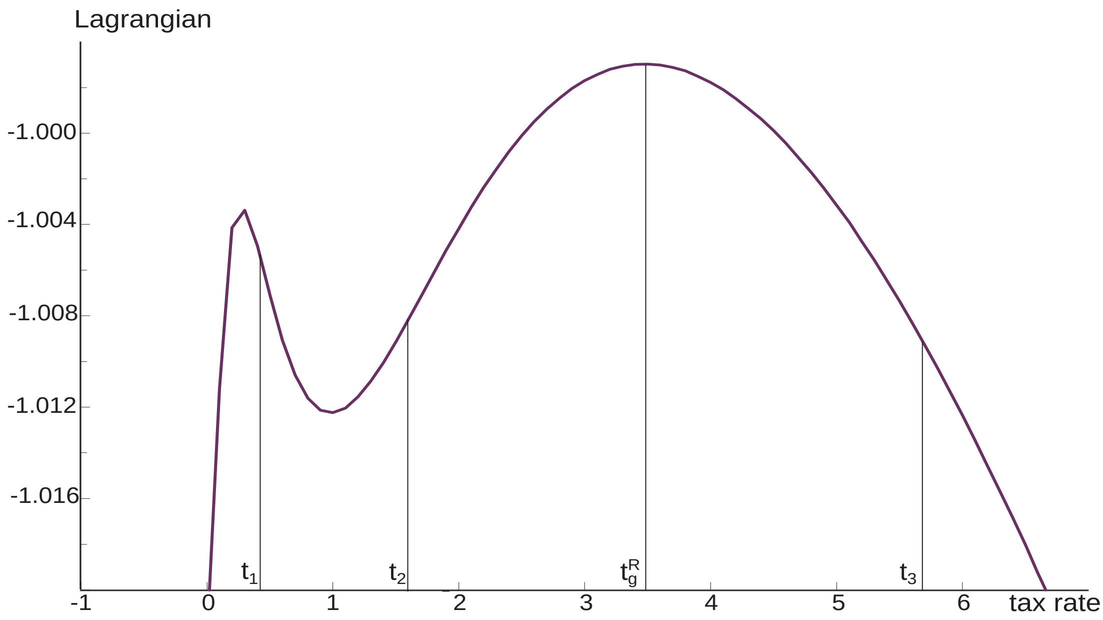

Figure 3: A Lagrangian function with two peaks

elasticities are identical across consumers, i.e., $\varepsilon_g(c) = \varepsilon_g$. Then, the contribution of commodity g to social welfare is

$$\mathcal{L}_g(t) = \begin{cases} A_g \dfrac{(1+t)^{-\varepsilon_g}}{\varepsilon_g - 1} \left[a_g(1+t) + \lambda t\left(\varepsilon_g - 1\right)\right] & \text{for } \varepsilon_g > 0,\ \varepsilon_g \neq 1 \\ A_g \left[a_g\left[-\ln\left(1+t\right) - 1\right] + \dfrac{\lambda t}{1+t}\right] & \text{for } \varepsilon_g = 1 \end{cases} \tag{12}$$

where $A_g = \int_C A_g(c) d\nu(c)$. For this specification, one can obtain a precise description of the optimal classification of goods in the different tax groups.

Under the assumptions of Section 6, one can derive the following results:

Lemma 3 *For any commodity g, with price elasticity $\varepsilon > 0$ and distributional characteristic a, the inequality $\mathcal{L}_g(t') > \mathcal{L}_g(t)$ is equivalent to*

$$\frac{a}{\lambda} < \phi\left(\varepsilon, t, t'\right)$$

where

$$\phi\left(\varepsilon, t, t'\right) = (1-\varepsilon)\left[1 + \frac{1}{\sqrt{(1+t')\,(1+t)}} \frac{\sinh\left(r\frac{\varepsilon}{2}\right)}{\sinh\left(r\frac{1-\varepsilon}{2}\right)}\right], \textit{ and } r = \ln\left(\frac{1+t'}{1+t}\right).$$

Proof: Using $t'(1+t')^{-\varepsilon} = (1+t')^{1-\varepsilon} - (1+t')^{-\varepsilon}$, the inequality $\mathcal{L}_g(t') > \mathcal{L}_g(t)$ rewrites

$$\left[(1+t')^{1-\varepsilon} - (1+t)^{1-\varepsilon}\right]\left(\frac{a}{\varepsilon - 1} + \lambda\right)\frac{1}{\lambda} > (1+t')^{-\varepsilon} - (1+t)^{-\varepsilon}.$$

Note that, for any real number σ,

$$(1+t')^{\sigma} - (1+t)^{\sigma} = 2\,(1+t)^{\sigma/2}\,(1+t')^{\sigma/2}\sinh\left(\frac{\sigma r}{2}\right)$$

where r is as defined in the Lemma. Thus, we get that $\mathcal{L}_g(t') > \mathcal{L}_g(t)$ is equivalent to

$$[(1+t')\,(1+t)]^{\frac{1}{2}}\sinh\left(r\frac{1-\varepsilon}{2}\right)\left(\frac{a}{\varepsilon-1}+\lambda\right)\frac{1}{\lambda}+\sinh\left(r\frac{\varepsilon}{2}\right) > 0.$$

Since $\sinh\left(r\frac{1-\varepsilon}{2}\right)$ has the same sign as $1-\varepsilon$, the last inequality rewrites

$$\frac{a}{\lambda} < \phi\,(\varepsilon,t,t')$$

where

$$\phi\,(\varepsilon,t,t') = (1-\varepsilon)\left[1+\frac{1}{\sqrt{(1+t')\,(1+t)}}\frac{\sinh\left(r\frac{\varepsilon}{2}\right)}{\sinh\left(r\frac{1-\varepsilon}{2}\right)}\right].$$

In the particular case $\varepsilon = 1$, it is sufficient to show that $\frac{a}{\lambda} < \phi\,(1,t,t')$ is equivalent to $\mathcal{L}_g(t') > \mathcal{L}_g(t)$ for $\varepsilon = 1$. Since $\sinh x$ is equivalent to x in the neighborhood of $x = 0$, we have

$$\phi\,(1,t,t') = \frac{1}{\sqrt{(1+t')\,(1+t)}}\frac{\sinh\left(\frac{r}{2}\right)}{\frac{r}{2}} = \frac{(1+t)^{-1}-(1+t')^{-1}}{\ln(1+t')-\ln(1+t)}$$

and $\mathcal{L}_g(t') > \mathcal{L}_g(t)$ is equivalent to

$$\frac{\lambda t'}{(1+t')} - a\ln(1+t') > \frac{\lambda t}{(1+t)} - a\ln(1+t)$$

$$\Leftrightarrow \lambda - \frac{\lambda}{(1+t')} - a\ln(1+t') > \lambda - \frac{\lambda}{(1+t)} - a\ln(1+t)$$

$$\Leftrightarrow \lambda\left[\frac{1}{(1+t)} - \frac{1}{(1+t')}\right] > a\left[\ln(1+t')-\ln(1+t)\right]$$

∎

Lemma 4 *For $t' > t$, the function $\phi(\varepsilon,t,t')$ is convex in its first argument.*

1. *Its slope at the origin is*

$$\frac{\partial\phi}{\partial\varepsilon} = \frac{1}{t'-t}\left[\ln\left(\frac{1+t'}{1+t}\right)-(t'-t)\right].$$

2. *When ε goes to ∞, ϕ is equivalent to*

$$(1-\varepsilon)\frac{t}{1+t}.$$

Proof: 1) Using the identity

$$\sinh a \cosh b + \sinh b \cosh a = \sinh(a+b),$$

a direct computation yields

$$\frac{\partial\phi}{\partial\varepsilon} = -1 - \frac{1}{\sqrt{(1+t')(1+t)}[\sinh{(1-\varepsilon)r/2}]^2}\left[\sinh\frac{(1-\varepsilon)r}{2}\sinh\frac{\varepsilon r}{2} - \frac{(1-\varepsilon)r}{2}\sinh\frac{r}{2}\right].$$

The desired formula follows when $\varepsilon = 0$, using the equality $\sqrt{(1+t')(1+t)}\sinh r/2 = (t'-t)/2$.

2) One can rewrite

$$\phi(\varepsilon, t, t') = (1-\varepsilon)\left[1 - \frac{1}{1+t}\frac{1-\exp(-\varepsilon r)}{1-\exp[(1-\varepsilon)r]}\right].$$

When ε goes to infinity, the result follows.

We finally show the convexity of ϕ with respect to ε, which is derived from the fact that its second derivative is positive. Indeed differentiating the expression obtained in 1) for the first derivative gives:

$$\frac{\partial^2\phi}{\partial\varepsilon^2} = \frac{r}{\sqrt{(1+t')(1+t)}}\frac{\sinh r/2}{[\sinh r(1-\varepsilon)/2]^2}\left[-1 + \frac{(1-\varepsilon)r}{2}\frac{\cosh\frac{(1-\varepsilon)r}{2}}{\sinh\frac{(1-\varepsilon)r}{2}}\right].$$

It is positive since $(x/\tanh x)$ is larger than 1 for all x (the tanh curve is below the 45^0 line for positive x, and above for negative x). ∎

B First-order condition of the planner problem without separability of individual preferences

For a (small) group of goods $\{g\} + dG$ around g, define $\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, s)$ to be the set of tax rates $\mathbf{t}'$ such that $t'_\gamma = t_\gamma$ for γ not in $\{g\} + dG$ and $t'_\gamma = s$ for γ in $\{g\} + dG$. Adapting Theorem 1, a necessary condition for the optimality of a partition $\mathbf{G}$ associated with tax rates $\mathbf{t}$ is that, for all k, for all g and all small enough dG such that $\{g\} + dG$ is in G_k, and for all h

$$\mathcal{L}[\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, t_k)] \geq \mathcal{L}[\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, t_h)].$$

We prove formula (8) of the text. The first-order condition for an interior maximum is

$$
\begin{aligned}
0 = &- \int_C \alpha_c \rho_c \int_{\{g\}+dG} \xi_\ell[\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, s), L_c, Y_c, c]\mu(\ell) d\ell \; d\nu(c) \\
&+ \lambda \int_C \int_{\{g\}+dG} \xi_\ell[\boldsymbol{\tau}_{g+dG}(\mathbf{t}, s), L_c, Y_c, c]\mu(\ell) d\ell \; d\nu(c) \\
&+ \lambda \int_C \left[\int_G t_\ell \frac{\partial \xi_\ell[\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, s), L_c, Y_c, c]}{\partial s} \mu(\ell) d\ell \right] d\nu(c).
\end{aligned}
$$

We want to get the limit of the above expression when dG goes to zero, after division by the weight $\mu(dG) = \int_{dG} \mu(\ell) d\ell$. The two first terms, as well as the last one, are easily dealt with. Indeed, define, with some abuse of notation:

$$
\xi_g[\boldsymbol{\tau}_g(\mathbf{t}, s), L_c, Y_c, c] = \lim_{dG \to 0} \int_{\{g\}+dG} \frac{\xi_\ell[\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, s), L_c, Y_c, c]}{\mu(dG)} \mu(\ell) d\ell.
$$

Then, the two first terms tend to

$$
\{-a_g[\boldsymbol{\tau}_g(\mathbf{t}, s)] + \lambda\} \;\; X_g[\boldsymbol{\tau}_g(\mathbf{t}, s)].
$$

The third term needs some more care. When taking the limit, one must separate the own price effect from the substitution effect on other goods:

$$
\frac{\partial \xi_g[\boldsymbol{\tau}_g(\mathbf{t}, s), L_c, Y_c, c]}{\partial s} = \lim_{dG \to 0} \int_{\{g\}+dG} \frac{1}{\mu(dG)} \frac{\partial \xi_\ell[\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, s), L_c, Y_c, c]}{\partial s} \mu(\ell) d\ell,
$$

$$
\frac{\partial \xi_{G_k \setminus \{g\}}[\boldsymbol{\tau}_g(\mathbf{t}, s), L_c, Y_c, c]}{\partial s} = \lim_{dG \to 0} \int_{G_k \setminus \{\{g\}+dG\}} \frac{1}{\mu(dG)} \frac{\partial \xi_\ell[\boldsymbol{\tau}_{\{g\}+dG}(\mathbf{t}, s), L_c, Y_c, c]}{\partial s} \mu(\ell) d\ell.
$$

The former limit is the own price elasticity, while the latter is the average substitution effect on the commodities[16] in the set $G_k \setminus \{g\}$, which only exists when substitution between commodities is not too 'large'. Finally, summing up on agents, define

$$
\frac{\partial X_g[\boldsymbol{\tau}_g(\mathbf{t}, s)]}{\partial s} = \int_C \frac{\partial \xi_g[\boldsymbol{\tau}_g(\mathbf{t}, s), L_c, Y_c, c]}{\partial s} d\nu(c),
$$

and

$$
\frac{\partial X_{G_k \setminus \{g\}}[\boldsymbol{\tau}_g(\mathbf{t}, s)]}{\partial s} = \int_C \frac{\partial \xi_{G_k \setminus \{g\}}[\boldsymbol{\tau}_g(\mathbf{t}, s), L_c, Y_c, c]}{\partial s} d\nu(c).
$$

The equations (6) and (8) of the text follow with s standing for the optimal rate applied to group G_k or commodity g.

[16] We use the notation $G_k \setminus \{\{g\} + dG\}$ as a short hand for $G_k \setminus [G_k \cap \{\{g\} + dG\}]$. Note that since, by construction, $\{g\} + dG$ is contained in a single member of the partition, say G_h, all the $G_k \setminus \{\{g\} + dG\}$'s coincide with G_k, for all k different from h.

	B&C	M&F	D	T&C	F&V	OUF	F	FO	B	W&S	T	DF	HGS	AC	CC	P&D	PT	LG	B&N	LS
Bread & Cereals	**-0.391**	-0.085	-0.154	-0.02	-0.175	0.385	-0.441	0.812	-0.227	-0.551	0.151	0.232	0.203	-0.381	0.093	0.146	-0.005	-0.474	-0.489	0.371
Meat & Fish	-0.026	**-0.843**	0.038	-0.019	-0.107	0.072	-0.082	0.003	-0.029	-0.187	-0.05	-0.088	-0.173	0.084	-0.024	0.171	0.064	-0.066	-0.062	0.325
Dairy	-0.07	0.057	**-0.247**	0.073	-0.175	-0.206	0.012	-0.043	-0.353	0.142	-0.072	0.063	0.148	-0.044	-0.281	0.175	-0.047	-0.395	-0.101	0.367
Tea & Coffee	-0.045	-0.139	0.359	**-0.32**	-0.266	-0.037	-0.089	-0.322	0.312	-0.338	0.086	-0.034	0.487	-0.768	0.191	0.159	-0.149	-0.077	0.017	-0.029
Fruits and Vegetables	-0.088	-0.176	-0.192	-0.059	**-0.345**	-0.173	-0.04	0.275	0.081	-0.005	-0.045	-0.135	-0.159	-0.37	-0.107	0.146	0.088	0.061	-0.1	0.344
Other Untax. Foods	0.291	0.178	-0.34	-0.013	-0.26	**-0.381**	-0.22	0.4	0.756	0.058	-0.356	-0.403	-0.205	-0.52	-0.367	0.037	0.047	0.306	-0.096	0.086
Standard VAT Food	-0.299	-0.184	0.017	-0.027	-0.054	-0.198	**-1.68**	0.486	0.354	-0.338	-0.221	-0.554	0.274	1.615	-0.016	0.052	-0.036	-1.11	0.118	0.8
Food out	0.356	0.004	-0.041	-0.063	0.24	0.232	0.314	**-0.19**	-0.674	0.044	0.336	-0.143	-0.875	-0.517	0.338	-0.181	-0.085	0.307	0.111	-0.514
Beer	-0.137	-0.058	-0.468	0.084	0.098	0.606	0.317	-0.932	**-1.408**	-0.077	0.388	0.883	0.244	0.285	0.489	-0.406	0.111	-0.3	-0.292	-0.427
Wine & Spirits	-0.637	-0.711	0.358	-0.174	-0.011	0.089	-0.576	0.116	-0.146	**-3.024**	0.603	0.063	2.455	1.975	0.877	0.3	-0.31	-0.599	-0.341	-1.307
Tobacco	0.084	-0.092	-0.088	0.021	-0.05	-0.262	-0.181	0.427	0.356	0.29	**-0.864**	0.19	-0.763	-0.089	-0.144	-0.204	0.1	0.203	-0.065	0.132
Domestic Fuels	0.054	-0.067	0.032	-0.003	-0.063	-0.124	-0.19	-0.076	0.339	0.013	0.08	**-0.324**	-0.269	-0.817	-0.07	-0.042	0.098	-0.085	0.179	0.335
Household G & S	0.026	-0.073	0.042	0.028	-0.041	-0.035	0.052	-0.257	0.052	0.274	-0.177	-0.149	**-0.892**	-0.074	-0.037	-0.199	0.007	0.044	0.127	0.281
Adult Clothing	-0.122	0.089	-0.031	-0.11	-0.237	-0.221	0.764	-0.378	0.151	0.548	-0.051	-1.127	-0.184	**-1.07**	0.01	0.111	0.083	1.213	0.141	-0.579
Children's Clothing	0.236	-0.204	-1.558	0.215	-0.539	-1.23	-0.06	1.953	2.045	1.923	-0.656	-0.76	-0.719	0.077	**-0.092**	0.148	0.538	-0.047	-0.678	-1.592
Petrol and Diesel	0.073	0.282	0.191	0.036	0.146	0.025	0.039	-0.207	-0.336	0.13	-0.184	-0.09	-0.771	0.174	0.029	**-0.383**	-0.083	0.124	-0.095	-0.1
Public Transport	-0.006	0.259	-0.125	-0.081	0.215	0.076	-0.065	-0.237	0.223	-0.329	0.219	0.514	0.07	0.318	0.26	-0.202	**-0.736**	-0.254	-0.691	-0.428
Leisure Goods	-0.241	-0.11	-0.441	-0.017	0.062	0.206	-0.834	0.357	-0.252	-0.264	0.186	-0.186	0.174	1.926	-0.009	0.126	-0.106	**-1.341**	-0.148	-0.089
Books & Newspapers	-0.462	-0.193	-0.21	0.007	-0.188	-0.12	0.165	0.239	-0.455	-0.279	-0.11	0.728	0.933	0.414	-0.253	-0.179	-0.534	-0.274	**-0.127**	-0.103
Leisure Services	0.087	0.251	0.188	-0.003	0.161	0.027	0.277	-0.276	-0.166	-0.266	0.056	0.339	0.512	-0.424	-0.148	-0.047	-0.082	-0.041	-0.025	**-1.421**

I. Elasticities of the demand for goods (rows) with respect to prices (columns)

	Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10	Overall	Av. Tax rate
Bread & Cereals	0.03877	0.03093	0.0274	0.02515	0.02306	0.02102	0.01902	0.01702	0.01461	0.01037	0.02273	0%
Meat & Fish	0.11791	0.10042	0.09222	0.08495	0.0778	0.07142	0.06378	0.05618	0.04764	0.03509	0.07474	0%
Dairy	0.08713	0.06931	0.06137	0.05619	0.05036	0.04583	0.04071	0.03565	0.03007	0.02107	0.04977	0%
Tea & Coffee	0.01854	0.01459	0.01245	0.01129	0.0101	0.00911	0.00805	0.00708	0.00587	0.0041	0.01012	0%
Fruits and Vegetables	0.06246	0.05681	0.05306	0.04979	0.04704	0.04451	0.04137	0.0382	0.03413	0.02606	0.04534	0%
Other Untax. Foods	0.04277	0.03724	0.03393	0.03238	0.03079	0.02921	0.02742	0.02586	0.02335	0.01817	0.03011	0%
Standard VAT Food	0.04098	0.03877	0.03733	0.03551	0.03464	0.03366	0.03226	0.0308	0.02864	0.02279	0.03354	17.50%
Food out	0.02647	0.03441	0.03994	0.04437	0.05048	0.05489	0.06071	0.06613	0.07147	0.06985	0.05187	17.50%
Beer	0.02131	0.03223	0.03664	0.03962	0.04116	0.0421	0.04287	0.04237	0.04194	0.03509	0.03753	28%
Wine & Spirits	0.0108	0.01427	0.01685	0.01828	0.01994	0.02109	0.02252	0.02337	0.02487	0.02487	0.01969	55.80%
Tobacco	0.04183	0.05082	0.05255	0.05075	0.04797	0.04394	0.03996	0.03516	0.02802	0.01803	0.0409	79.10%
Domestic Fuels	0.16981	0.14192	0.12375	0.11068	0.0996	0.08992	0.07856	0.0681	0.05654	0.03936	0.09782	5%
Household G & S	0.13618	0.15314	0.16222	0.17019	0.17491	0.1814	0.18828	0.1958	0.19979	0.20161	0.17635	17.50%
Adult Clothing	0.03831	0.05315	0.06014	0.06612	0.07077	0.07433	0.07939	0.08317	0.08745	0.09643	0.07093	17.50%
Children's Clothing	0.00202	0.00533	0.00752	0.00898	0.01	0.01058	0.01102	0.01164	0.01177	0.01097	0.00898	0%
Petrol and Diesel	0.01385	0.02871	0.03672	0.04445	0.04988	0.05418	0.05819	0.05899	0.05909	0.04989	0.04539	73.25%
Public Transport	0.02006	0.02115	0.02119	0.02007	0.01939	0.01912	0.0178	0.01697	0.01542	0.01481	0.0186	17.50%
Leisure Goods	0.01523	0.02362	0.03047	0.03485	0.04046	0.04615	0.05027	0.05718	0.06735	0.08105	0.04466	17.50%
Books & Newspapers	0.03357	0.0297	0.02724	0.02514	0.0245	0.02337	0.02161	0.02042	0.01878	0.01633	0.02407	0%
Leisure Services	0.06199	0.06348	0.06704	0.07124	0.07714	0.08419	0.09621	0.10993	0.13318	0.20406	0.09685	17.50%

II. Budget shares by consumption deciles and aggregate; indirect VAT and excise tax rate (fraction of consumption price)