Optimal Indirect Taxation and Income Inequality

Corrado Benassi[*] and Emanuela Randon[†]

This version: June 2018

Abstract

Policy responses to growing inequality are a priority issue in the agenda of many a government. We provide a general framework to analyse the interaction between income distribution and indirect taxation, in order to assess how the former shapes optimal indirect taxes and the optimal responses of such taxes to inequality changes. We find policy prescriptions at the aggregate level, as well as for the single tax adjustment, that depend on the structural progressivity of the indirect tax system. An increase in inequality not countered by any optimal tax adjustment amounts to an implicit taxation, due to changes in the tax burden distribution driven by the impact of the distributional shock on individual consumption patterns. We find conditions for a simple formulation of optimal taxation policies useful for empirical work.

Keywords: Optimal Commodity Taxation; Income Distribution.
JEL classification: H21; D63; D11.

1 Introduction

Income and wealth inequality have been rising in many countries over the last decades. While causes and effects of inequality and its impact on the

[*]Department of Economics - University of Bologna. Piazza Scaravilli 1, 40126 Bologna (Italy). Phone:+390512098667; Fax:+390512098040. Email: corrado.benassi@unibo.it.
[†]Corresponding Author: Department of Economics - University of Bologna. Strada Maggiore 45, 40125 Bologna. Phone:+390512092625; Fax:+390512098143. Email: emanuela.randon@unibo.it.

welfare and behaviour of people are always likely to generate controversy (e.g., Piketty and Saez, 2014), policy responses to growing inequality have become a priority issue in the agenda of many a government. Among the various questions involved in such an agenda, that of how taxation should react to inequality ranks quite high.

Against such a background, it is remarkable that indirect taxation is becoming quite relevant: urgent advices call for governments to decrease direct taxation and shift the tax burden onto more "growth-friendly" indirect taxes,[1] as OECD (2014), the European Commission (2015) or the IMF keep suggesting a more intense use of indirect taxation – more generally, tax shifting (from labour to consumption taxes), is often becoming a policy priority.[2] On the other hand, it is not unusual that a 'flat rate', distributionally neutral income tax be advocated in public debates:[3] so that, with income taxation losing any redistributive effect in the long run, any such effect would fall on other forms of taxation, and consumption taxes would easily play a key role.

All this brings to the fore the issue of the distributional features of indirect taxation,[4] which becomes all the more relevant as indirect taxation is one of the main sources of government revenue:[5] given the huge amount of

[1] Myles (2009a,b,c) provides an excellent survey of the linkages between growth and taxation.

[2] See for instance the recent concluding statements of IMF Article IV Mission for Netherlands (2015) or Belgium (2016).

[3] See, e.g., "Why the Flat Tax Is More Popular Than Ever?", Time Magazine, Amity Shlaes, 10 November 2015: once adopted, a switch from a progressive to a proportional fiscal regime would immediately increase income inequality.

[4] This issue has a long history: "Historically, the role of indirect taxes has often been redistributional: for example, in seventeenth-century England the taxation of silks, coffee, and newspapers may well have been more progressive than a one-shilling-a-head poll tax" (Atkinson and Stiglitz, 1980, p.432, f.note 3). For a recent survey, see e.g. Burgert and Roeger (2014), or European Commission (2015).

[5] According to OECD (2015), taxation of goods and services in 2013 accounted for 32.7% of the tax revenue in OECD countries (11% of GDP), and 17.4% of the US tax yield (4.4% of GDP). The most important form of indirect taxation in Europe, i.e. VAT, collected 976.9 billion euros in 2014 among the EU-27 member States, despite the 158.5 billion tax flow

financial resources absorbed by such taxes, the design and implementation of an optimal indirect taxation system *vis à vis* the distributional features of the economy is an important choice for the public sector, as well as for the wellbeing of a society. Given these premises, this paper focuses on the interplay between income distribution and indirect taxation, in order to evaluate how income distribution affects second-best optimal commodity taxes and the optimal response of such taxes to inequality changes.[6]

Our modeling strategy rests on two premises. First, we treat income y as exogenous: our setting is equivalent to a scenario in which the numeraire commodity (leisure) is untaxed, and y is the individual time endowment (plus non-earned resources), i.e. wealth.[7] An alternative interpretation would have our framework suggesting a way to look at the interplay of direct and indirect taxation, whenever the distributional shift variable is seen as a proxy for the progressivity of (non optimal) direct taxation on wealth.[8] Sec-

lost in informal or black markets.

[6]Optimal commodity taxation has been strictly related to second-best policies since the seminal paper by Ramsey (1927), who simply ruled out the possibility of using lump sum taxes in a single agent economy. Since then, the main theoretical body has been developed in the 70s and 80s (e.g.: Diamond and Mirrless, 1971), the focus being on the features of "second-best" efficient and fair taxes, under a number of different conditions. Diamond 's covariance (1975) was the first simple formulation of second best optimal taxes. Boadway (2012) provides a recent survey.

[7]In principle, leisure could be taxed and some other commodity used as a numéraire. An alternative formulation would follow Salanie (e.g., 2003, p.111) and assume that $v(p, y_A)$ is the indirect utility obtained from maximising h's direct utility $U^h(q_i, \ell)$ subject to the budget constraint $\sum_i (1 + t'_i)q_i = (1 - \tau)\ell + y_A$, where ℓ is the supply of labor, t'_i is the linear tax on goods, τ is the linear tax on wages, $p = 1 + t$ and $y_A = y - A$, A the lump sum tax: within this framework, one can parameterize the indirect taxes so that the tax system (t'_i, τ) is equivalent to the tax system (t_i, 0).

[8]The government may set non optimal taxes due to political or institutional constraints (leading, e.g., to trading off efficient direct taxation against better targetability: Castanheira *et al.*, 2012). At the theoretical level, the issue of optimal taxation of commodity *vs* (endogenous) income is standardly addressed using the AS (Atkinson and Stiglitz, 1976) framework, which spurred a great deal of research on second best interventions (a general assessment is provided by Boadway and Pestieau, 2003, and Boadway 2012, pp.57-85). Importantly in our perspective, Laroque (2005) and Kaplow (2006) show that the AS results can be extended to the case where non linear income taxation is non optimal (see also

ondly, we model inequality shocks to the income distribution as second-order stochastic-dominance shifts, such that a change of an exogenous parameter of the income density function standardly identifies an increase in inequality for a given mean.

We first present in the next section our basic framework, which allows to consider how income distribution affects the "many-person Ramsey tax rule" and Diamond's covariance (Diamond, 1975) between the social marginal utility of income and the individual quantity shares. It is generally assumed that once excise taxes are coupled with poll taxes, this covariance is positive for necessities and negative for luxuries – this assumption relying on the supposed correlation between the individual consumption shares and the social marginal utility of income, due to an inequality adverse government assigning larger social weight to the consumption of the poor. To our knowledge, however, no argument is to be found connecting the sign of the covariance to more intrinsic features, such as the price or income elasticity of market demand. Here we show that with a utilitarian social planner, the covariance can be explicitly connected to the income elasticity of individual demand and is always negative for normal commodities, leading in this case to a nonnegative discouragement index. We also show how linearization can provide a simple covariance formula which may in principle be useful for empirical estimation.

We then assess the impact of inequality changes on optimal indirect taxation. In this perspective, we find that taxation policy should take into

Piketty and Saez, 2013); however, one might argue that the AS approach makes it quite difficult to obtain clear-cut results without crucial restrictions on utility (i.e. weak separability), which have been questioned on empirical grounds (e.g., Browning and Meghir, 1991; Pirttilä and Suoniemi, 2014). This being so, one can arguably look at the impact of distributive shocks on optimal commodity taxes, independently of the source of such shocks – if they are due to (non optimal) direct taxation, one can in principle assess how (non linear) changes in direct taxation affect optimal commodity taxation.

account the structural regressivity (progressivity) features of the commodity tax system, and we derive conditions for the emergence of a welfare-efficiency tradeoff connecting that degree of progressivity with the type of shock which hits the income distribution. We also assess the "implicit tax" of higher inequality, which emerges when the government does not adjust commodity taxation to a more unequal income distribution. This "implicit tax" is given by changes in the distribution of the tax burden across commodities, driven by income distribution changes through their effects on demand quantities.

The paper is organised as follows. In Section 2 we analyse how income distribution affects the optimal indirect taxes. In Section 3, we assess the impact of inequality changes on such taxes. We discuss the welfare-efficiency tradeoff, and consider the implicit tax of inequality. Concluding remarks are gathered in Section 4.

2 The basic framework

We consider consumers' heterogeneity as solely due to income differences. Income $y \in \mathcal{Y}$ is continuously distributed over some positive support $[y_{\min}, y_{\max}]$ $= \mathcal{Y}$, with $F : \mathcal{Y} \times \Theta \to [0, 1]$ the associated distribution. $\Theta \subset \mathbb{R}$ is some parameter space: $\theta \in \Theta$ is a distribution parameter which – as will be clear below – we use to measure inequality. We denote the income density by $f(y, \theta) = \frac{\partial F}{\partial y}$, so that $\mu = \int_{\mathcal{Y}} y f(y, \theta) dy > 0$ is aggregate (mean) income.

The government chooses n second-best optimal commodity tax rates $t = (t_1, ..., t_n)$ and a uniform poll tax $A < y_{\min}$, by maximising a social welfare function given the revenue budget constraint. As the government knows the income distribution and the traded quantities, but not the identity of the single consumer, we rule out the possibility of using first order policy interventions such as personal lump sum taxes. We define the social welfare function as

$$V(p, A; \theta) = \int_{\mathcal{Y}} v(p, y_A) f(y, \theta) dy, \tag{1}$$

where $v(p, y_A)$ is the indirect utility of an individual with net income $y_A = y - A$, facing the vector of $n+1$ gross prices $p = (p_0, p_1, ..., p_n)$ such that $p = \widetilde{p} + t$, with $\widetilde{p}$ the given vector of net prices and $p_0 = \widetilde{p}_0 = 1$. Accordingly, the consumer's budget constraint is $\sum_i p_i q_i + q_0 = y$,[9] while the government revenue constraint is

$$R = \sum_i t_i Q_i(p, A; \theta) + A \tag{2}$$

where R is the exogenous revenue target, $Q_i(p, A; \theta) = \int_{\mathcal{Y}} q_i(p, y_A) f(y, \theta) dy$ and $q_i(p, y_A)$ are respectively the aggregate and the individual demand for commodity i.

This framework is obviously consistent with the classical Diamond (1975) formulation. On the one hand, the social marginal utility of income is

$$\gamma(t, y) = \frac{\partial v}{\partial y} + \lambda \frac{\partial \tau}{\partial y} \tag{3}$$

where $\tau = \sum_i t_i q_i(p, y_A)$ is the individual commodity tax burden.[10] On the other hand, working out the first order conditions of the standard welfare problem given by maximizing (1) subject to (2), the constraint multiplier λ appearing in (3) equals the expected value of social marginal utility, i.e.

$$\lambda = E(\gamma) = \int_{\mathcal{Y}} \gamma(t, y) f(y, \theta) dy. \tag{4}$$

[9] In the summation operator $\sum_i$ the index i is supposed to run from 1 to n.

[10] In Diamond's (discrete) formulation, the social marginal utility of income would be

$$\gamma(t, y) = \frac{\partial W}{\partial v^h} \frac{\partial v^h}{\partial y^h} + \lambda \frac{\partial \tau}{\partial y^h}$$

In our case, where the social welfare function W is strictly utilitarian, the social planner's higher weight for the poor would obviously be the result of the marginal utility of income being decreasing – an assumption we shall make in the following sections (apparently supported by the empirical evidence: eg, Layard *et al.*, 2008).

One can also show that the covariance between the social marginal utility of income and the individual quantity shares can be written as

$$\sigma(\gamma, q_i/Q_i) = \int_{\mathcal{Y}} \gamma\,(y,t)\,[\varphi_i(y,p;\theta) - f(y;\theta)]\,dy \tag{5}$$

where $\varphi_i(y,p;\theta) = q_i(p,y_A)f(y,\theta)/Q_i$ is the density which gives the distribution of demand for commodity i by income. The percentage change in quantities associated with optimal taxation will be

$$\frac{dQ_i^h}{Q_i} = \tfrac{1}{\lambda}\sigma(\gamma, q_i/Q_i) = \overline{\sigma}(\gamma, q_i/Q_i) \tag{6}$$

where $dQ_i^h = \int_{\mathcal{Y}} \left(\sum_j t_j \frac{\partial q_j^h}{\partial p_i}\right) f(y,\theta)dy$, q_j^h is compensated individual demand, and $\overline{\sigma}$ denotes the mean-weighted covariance.[11]

Let now ε_i denote the income elasticity of the demand for commodity i. We can then use in our framework the definition provided by Salanie (e.g., 2003, p.68) of the "discouragement index" for commodity i, i.e. $D^i = -\overline{\sigma}(\gamma, q_i/Q_i)$, in order to cast as follows our first result:

Proposition 1 *If the marginal social utility of income is decreasing in income: (a) the covariance between the social marginal utility of income and the individual quantity shares is non-positive (negative) for all normal (inferior) goods, implying a non-negative (positive) discouragement index for all such goods; (b) for any two commodities i and j, if $\varepsilon_i\,(y,p) > \varepsilon_j\,(y,p)$ for all $y \in \mathcal{Y}$, then $D^i > D^j$, i.e. commodity i should be taxed more heavily than commodity j.*

Proof. See Appendix B. ■

If the marginal social utility of income is decreasing in income, it is generally supposed that the covariance $\sigma(\gamma, q_i/Q_i)$ is positive for necessities and

[11] Equations (4) and (5), and the resulting equation (6), are derived in Appendix A.

negative for luxuries, as a consequence of how the social marginal utility of income is assumed to behave with respect to the individual consumption shares. By contrast, Proposition 1 suggests that, once we link the Diamond covariance to the income distribution and $\gamma(t, y)$ is decreasing, covariance is always negative (positive) for normal (inferior) goods. In the important case of normal commodities, the implication is that optimal commodity taxation lowers consumption: the simple intuition behind this is that if commodities are normal, individual demand is increasing in income and so its covariance with a decreasing social marginal utility is bound to be negative. In such a case, optimal taxation enforces a decrease in consumption that however should be heavier for goods whose income elasticity is higher, i.e. luxuries. In a way, this result goes in the same direction as the policy prescriptions generally associated to covariance plus lump-sum models, according to which taxation should "encourage" the consumption of necessities and "discourage" that of luxuries.

The results of Propositions 1 can be further developed if one is willing to assume that $\gamma(t, y)$ can be approximated linearly, i.e. that the derivative $\partial\gamma/\partial y$, α say, is independent of income (this implies nil third order derivatives of utility and the tax burden with respect to income). Let $\mu_i = \int_{\mathcal{Y}} y\varphi_i(y, p; \theta)dy$ be the consumption-weighted average income; it is then immediate that

Corollary 1 *If the social marginal utility is linear in income, then*

$$\sigma(\gamma, q_i/Q_i) = \alpha(\mu_i - \mu), \tag{7}$$

where μ_i is the consumption-weighted average income and μ is average income, such that if commodities are normal $\mu_i - \mu > 0$.

Proof. Equation (7) follows trivially from integrating by parts (5), and the fact that for any continuous distribution F its mean satisfies $\mu = y_M - \int_{\mathcal{Y}} F(y,\theta)dy$. Recalling our definition of $\varphi_i(y,p;\theta) = q_i(p,y)f(y,\theta)/Q_i$, $\mu_i = \int_{\mathcal{Y}} y\varphi_i(y,p;\theta)dy$: since, as shown in the proof of Proposition 1, the distribution Φ_i dominates F in the first order sense when i is normal, it follows that $\mu_i > \mu$ in this case. ■

This formula may be useful for empirical purposes, as optimal indirect taxation is connected in a simple way to the difference between μ_i and μ, the sign of which depends on commodity i being normal or inferior.[12] Under linearity of γ, there results a clear cut relationship between the discouragement index (and hence the direction of taxation) and the difference between the consumption-weighted average income and mean income. Indeed, (7) implies that

$$\frac{D_i}{D_j} = \frac{\mu_i - \mu}{\mu_j - \mu}$$

so that taxation is heavier for commodity i than commodity j whenever $\mu_i > \mu_j$.

3 Results and Discussion

In the previous section we have presented our basic framework and the way second-best optimal indirect taxation can be connected to income distribution. In this section we consider how a change in income distribution, and in particular a change in the degree of inequality, affects optimal commodity taxation. To do so we need both a precise measure of inequality, and a convenient way of formalizing the income progressivity of the tax system.

[12]As μ_i is a weighted average of income, where the weights are given by the shares of consumption of commodity i by consumers whose income is y, it is easily seen that $\mu_i - \mu \gtrless 0$ as commodity i is normal or inferior.

As to the former, we identify inequality by the standard notion of second order stochastic dominance: i.e., we assume that an increase in θ signals an increase in inequality for given mean μ, as we take θ to be a (inverse) parameter of second order stochastic dominance.[13] As to progressivity, we start from the simple observation that the interplay between income levels and the burden of taxation can be assessed by referring to the income concavity or convexity of τ. If the latter is convex in income, the (net) income elasticity of τ will be higher than one, and the ratio τ/y_A will be increasing in income. In such a case, the indirect tax structure can be deemed "progressive".[14] On the other hand, for any given vector t of tax rates, the indirect tax burden τ of an individual with net income $y_A = y - A$ will be convex or concave, according as luxuries or necessities prevail in his budget set. If luxuries (necessities) dominate, such that $\partial^2\tau/\partial y^2 > (<)0$, then the system is progressive (regressive). Indeed, if the former (say) is the case, a marginal increase in income is bound to imply a more than proportional increase in the indirect taxes paid out.[15]

Under this stipulation, in this section we first consider the impact of an inequality increase on optimal tax rates and then the related implications

[13] An increase in θ signals an increase of inequality in the second order sense if $\int_{y_m}^{y} (\partial F/\partial\theta)\, dz \geq 0$ for all $y \in \mathcal{Y}$. As is well known, under this definition lower θ means a less unequal distribution – an inequality averse social planner would prefer it to a higher θ distribution; if the further restriction is added that $\int_{y_m}^{y_M} (\partial F/\partial\theta)\, dz = 0$ (i.e., mean income is not altered by changes in θ), θ ranks equal mean distributions by their Lorenz curve (Atkinson, 1970). See Lambert (2001, ch.3) for an overall assessment of the welfare-theoretic foundations of inequality measures.

[14] One should also add the boundary condition that $\partial\tau/\partial y \geq \tau/y$ for $y - A = y_{\min}$. More generally, convexity of $\tau(t, y)$ wrt y implies progressivity, i.e. that $\tau(t, y)/y$ is increasing in y , if $\tau(t, 0) = 0$ (e.g., Lambert, 2001, p.193). Clearly income concavity of τ would imply that the ratio τ/y_A is decreasing in income, and in this sense the tax vector t may be called "regressive".

[15] We a slight *abus de language* we identify luxuries (necessities) with convexity (concavity) of the relevant Engel curves, which boils down to the standard elasticity definition if such curves start from the origin. We assume that the sign of $\partial^2 q_i/\partial y^2$ does not change with y.

in terms of aggregate welfare. Our results on higher inequality as implicit taxation are gathered in the final part of the section.

3.1 Optimal taxation and inequality

If overall taxation, as opposed to the target tax revenue R, is defined as $T(t, A; \theta) = \sum_i t_i Q_i(p, A; \theta) + A$, the revenue constraint (2) boils down to $R = T(t, A; \theta)$, implicit differentiation of which yields

$$\sum_i \frac{\partial T}{\partial t_i}\frac{dt_i}{d\theta} + \frac{\partial T}{\partial A}\frac{dA}{d\theta} = -\frac{\partial T}{\partial \theta}. \tag{8}$$

Under optimality, a shock to the income distribution should be such that the optimality conditions still hold (Appendix A, eqs (A.1a,b)). Thus (8) becomes

$$\sum_i \beta_i \frac{dt_i}{d\theta} + \frac{dA}{d\theta} = -\frac{\partial T}{\partial \theta}\frac{1}{1-\overline{\tau}} \tag{9}$$

where $\overline{\tau} = \int_{\mathcal{Y}} \frac{\partial \tau}{\partial y} f(y, \theta) dy < 1$ is average marginal taxation wrt income,[16] and the specification of the weights β_i can be recovered from the same optimality conditions.[17] Equation (9) sheds some light on the effects of an increase in inequality on optimal taxation:

Proposition 2 *Assume $t_i > 0$ for some i, and let an increase in θ signal an increase in inequality for given mean. Then, if indirect taxation is progressive*

[16]Let $\overline{v} = \int_{\mathcal{Y}} \frac{\partial v}{\partial y} f(y, \theta) dy > 0$ be average marginal utility. By substituting in the definitions of γ and λ given in (3), one has $\overline{\tau} = \int_{\mathcal{Y}} \frac{\partial \tau}{\partial y} f(y, \theta) dy = 1 - \overline{v}/\lambda$ and hence $1 - \overline{\tau} = \overline{v}/\lambda > 0$. Notice that since $1 - \overline{\tau} = \int_{\mathcal{Y}} g(\widetilde{p}, t; y) f(y, \theta) dy$, with $g(\widetilde{p}, t; y) = \frac{\partial q_0(\widetilde{p}+t, y)}{\partial y} + \sum_i \widetilde{p}_i \frac{\partial q_i(\widetilde{p}+t, y)}{\partial y}$ $(g(\widetilde{p}, 0; y) = 1)$, at the optimum "normal behaviour" dominates in the sense that $g(\widetilde{p}, t; y) > 0$. Since this evaluates income effects of the taxed commodities at their gross prices p_i but weighs them with net prices $\widetilde{p}_i$, $g(\widetilde{p}, t; y) > 0$ if all commodities are normal at p, though it is also consistent with a subset of commodities being inferior.

[17]From (A.1b) $\partial T/\partial A = \overline{v}/\lambda$, so that $\partial T/\partial A = 1 - \overline{\tau}$ (see f.note 17). On the other hand, $\int_{\mathcal{Y}} \frac{\partial v}{\partial y} q_i(p, y) f(y, \theta) dy = \sigma(\partial v/\partial y, q_i) + \overline{v} Q_i$, so that from (A.1a) $\partial T/\partial t_i = (1 - \overline{\tau})[\overline{\sigma}(\partial v/\partial y, q_i) + 1] Q_i$ and, multiplying both sides of (8) by $1/(1 - \overline{\tau}) > 1$ equation (9) follows. In particular, $\beta_i = [\overline{\sigma}(\partial v/\partial y, q_i) + 1] Q_i$.

(regressive), optimality requires an average decrease (increase) of the tax rates (t, A).

Proof. First notice that $1-\overline{\tau}>0$, so the sign of $\sum_i \beta_i \frac{dt_i}{d\theta}+\frac{dA}{d\theta}$ depends on $\frac{\partial T}{\partial \theta}$. Letting subscripts denote derivatives, the latter can be written as $\int_{\mathcal{Y}} \tau(t,y) f_\theta(y,\theta) dy$, and since an increase in θ amounts to a second order stochastic dominance shift of the distribution, it follows from the general properties of stochastic dominance that $\frac{\partial T}{\partial \theta}=\int_{\mathcal{Y}} \tau(t,y) f_\theta(y,\theta) dy>0$ if $\tau(t,y)$ is convex in y, so that the system is progressive: hence $\sum_i \beta_i \frac{dt_i}{d\theta}+\frac{dA}{d\theta}<0$. By the same token, if $\tau(t,y)$ is concave in y, $\sum_i \beta_i \frac{dt_i}{d\theta}+\frac{dA}{d\theta}>0$. ■

Proposition (2) and equation (9) lend themselves to two main comments. First, they point to the fact that the overall reaction of optimal taxation to an increase of inequality is driven by the concavity of convexity of the individual tax burden (i.e. $\tau=\sum_i t_i q_i(p, y_A)$). The standard features of stochastic dominance imply that an inequality increase would yield *per se* a *virtual* higher (lower) tax intake if the individual tax schedule is convex (concave). Hence, for a given revenue target R, it calls on average for lower (higher) tax rates (t, A) if the system if progressive (regressive).

Secondly, (9) allows to disentangle the commodity-specific from the economy-wide components in the optimal tax adjustment induced by an inequality change. The overall required tax adjustment $-(\partial T/\partial \theta)$ is weighted by a measure of the system's overall progressivity, to give the aggregate progressivity-adjusted tax change. The latter can in turn be seen as the weighted average of the tax adjustments required of the lump sum A and the single commodity rates t_i, where $\sum_i \beta_i \frac{dt_i}{d\theta}$ is the contribution coming from changing commodity taxation. The weights β_i's can be written out as:[18]

$$\beta_i = \left[\overline{\sigma}\left(\partial v/\partial y, q_i\right) + 1\right] Q_i. \tag{10}$$

If the marginal utility of income is decreasing, the weight β_i is larger, the lower the correlation (in absolute value) between individual demand q_i and the marginal utility of income $\frac{\partial v}{\partial y}$, and the smaller the market size Q_i. To fix ideas, suppose $Q_i = Q_j$ (markets are of equal size): then

$$\frac{\beta_i}{\beta_j} = \frac{\overline{\sigma}\left(\partial v/\partial y, q_i\right) + 1}{\overline{\sigma}\left(\partial v/\partial y, q_j\right) + 1}$$

so that, given that decreasing marginal utility and normal commodities imply $\overline{\sigma}(\cdot,\cdot) < 0$, the larger weight will be associated with the (normal) commodity with the lower marginal utility in absolute value.

One can derive a convenient expression for these weights, by observing that using definition (3) of γ, there arises a natural connection among covariances, such that

$$\sigma\left(\frac{\partial v}{\partial y}, \frac{q_i}{Q_i}\right) = \sigma\left(\gamma, \frac{q_i}{Q_i}\right) - \lambda\sigma\left(\frac{\partial \tau}{\partial y}, \frac{q_i}{Q_i}\right) \tag{11}$$

which may be relevant from an observational point of view if one adopts a linear approximation for γ. In this case equations (7) and (11) allow to write out

$$\sigma\left(\frac{\partial v}{\partial y}, \frac{q_i}{Q_i}\right) = \alpha\left(\mu_i - \mu\right) - \lambda\sigma\left(\frac{\partial \tau}{\partial y}, \frac{q_i}{Q_i}\right)$$

which connects the covariance between consumption and the marginal utility of income to variables which are in principle observable. The weights involved would then be

$$\begin{aligned}
\beta_i &= \left[\overline{\sigma}\left(\partial v/\partial y, q_i\right) + 1\right] Q_i \\
&= \left[\tfrac{1}{(1-\overline{\tau})\lambda}\sigma\left(\tfrac{\partial v}{\partial y}, \tfrac{q_i}{Q_i}\right) + 1\right] Q_i \\
&= \left[\tfrac{\alpha}{\lambda}\tfrac{\mu_i-\mu}{1-\overline{\tau}} - \tfrac{1}{1-\overline{\tau}}\sigma\left(\tfrac{\partial \tau}{\partial y}, \tfrac{q_i}{Q_i}\right) + 1\right] Q_i
\end{aligned}$$

[18]See f.note 15.

where one should recall that under linearity $\frac{\alpha}{\lambda}(\mu_i - \mu) = -dQ_i^h/Q_i$.[19]

Finally, a noteworthy special case of the above may be that where the government is able to adjust only one tax rate, say t_i:

Proposition 3 *Assume $t_i > 0$, let an increase in θ signal an increase in inequality for given mean, and let $\frac{dA}{d\theta} = 0 = \frac{dt_j}{d\theta}$ for $j \neq i$. Then, the optimal tax adjustment for a normal commodity i is such that if the system is progressive (regressive) the negative (positive) tax rate adjustment is larger, the lower the covariance between the demand of good i and the marginal utility of income.*

Proof. From (9), if $\frac{dt_j}{d\theta} = 0$ for $j \neq i$ then

$$\frac{dt_i}{d\theta} = -\frac{\partial T}{\partial \theta}\left[\overline{\sigma}\left(\partial v/\partial y, q_i\right) + 1\right]^{-1}\frac{Q_i}{1-\overline{\tau}}$$

Since $1 - \overline{\tau} > 0$ and $\overline{\sigma}(\partial v/\partial y, q_i) + 1 \geq 0$, $\frac{dt_i}{d\theta} > 0$ if $\frac{\partial T}{\partial \theta} < 0$, which occurs if the system is regressive, i.e. necessities weights more (less) than luxuries (clearly, $\frac{dt_i}{d\theta} > 0$ if the system is progressive). Since $\overline{\sigma}(\partial v/\partial y, q_i) < 0$, the adjustment is larger the lower $\overline{\sigma}$ in absolute value. ■

Consistently with Proposition 2, the sign of the adjustment depends on the aggregate features of the overall tax system (as $\partial T/\partial \theta$ will include the cross effects of a change of t_i on all commodities), while its strength will depend inversely on the (normalized) covariance. Two observations are called for in this respect. First, the focus on one tax rate allows Proposition 3 to give a precise content to the *adjustment* $dt_i/d\theta$, as opposed to the *weights* of the adjustments considered in Proposition 2. Secondly, one might argue that, e.g., the (negative) correlation between marginal utility and consumption is

[19] We use the fact that average marginal utility satisfies $\overline{v} = (1-\overline{\tau})\lambda$ at the optimum, while $\overline{\sigma}(\partial v/\partial y, q_i) = \sigma(\partial v/\partial y, q_i)/\overline{v}Q_i$, and $\sigma(\partial v/\partial y, q_i) = Q_i\sigma(\partial v/\partial y, q_i/Q_i)$.

likely to be weaker with respect to necessities than luxuries: in which case, the single-rate reaction to an inequality increase under (say) a progressive tax system, would involve the alternative between a "large" tax reduction on necessities and a "smaller" reduction on luxuries.[20] More generally, in a framework where only one commodity is targeted for fiscal adjustment, any information about the strength of its correlation with marginal utility should be taken into account when singling out what commodity should be involved.[21]

3.2 Optimal taxation, inequality, and welfare

Consider social welfare at the optimum:

$$W(R,\theta) = V(\widetilde{p} + t(R,\theta), A(R,\theta); \theta)$$

differentiation of which yields

$$\begin{aligned} \frac{dW}{d\theta} &= \sum_i \frac{\partial V}{\partial t_i}\frac{dt_i}{d\theta} + \frac{\partial V}{\partial A}\frac{dA}{d\theta} + \frac{\partial V}{\partial \theta} && (12) \\ &= \lambda \frac{\partial T}{\partial \theta} + \frac{\partial V}{\partial \theta} && (13) \end{aligned}$$

after substitution from the first order conditions from Appendix A and using (8).[22] This expression calls for two observations. The first one is that the direct effect of an inequality increase on social welfare is in general negative, as $\partial V/\partial\theta < 0$ whenever $v(p,y)$ is concave in y, which will be the case if the marginal utility of income is decreasing (or, in a different context, the

[20]Clearly, the opposite dilemma would present itself under a regressive tax system, which would lead to a "large" tax increase for luxuries and a "smaller" tax increase for necessities.

[21]For the sake of completeness, one should add that if the single tax adjustment takes place through the lump-sum component, it will have to be $\frac{dA}{d\theta} = -\frac{\partial T}{\partial \theta}\frac{1}{1-\overline{\tau}}$, depending unsurprisingly on the existing tax structure as summarized by $\overline{\tau}$.

[22]That is, using the fact that at a maximum $\partial V/\partial p_i = -\lambda \partial T/\partial t_i$ and $\partial V/\partial A = -\lambda \partial T/\partial A$.

social planner is inequality averse). This is the direct effect of inequality, *independent of the optimal tax adjustments*. Secondly, however, the latter can affect welfare in the same or in the opposite direction, depending on the progressivity of the tax system:

Proposition 4 *Suppose an increase in θ signals an increase in inequality for given mean income. Then (a) if the commodity taxation system is progressive there is a welfare gain in the optimal tax adjustments; (b) if the commodity taxation system is regressive, there is a welfare-efficiency tradeoff.*

Both results are a direct implication of Proposition 2. On the one hand, if the system is progressive, an increase in inequality leads to an average decrease in taxation ($\frac{\partial T}{\partial \theta} > 0$), so that if the tax system is sufficiently progressive ($\lambda \frac{\partial T}{\partial \theta} > \frac{\partial V}{\partial \theta}$) there will be a net increase in welfare.[23] On the other hand, if necessities weigh more in the individual tax liability than luxuries, then $\lambda \frac{\partial T}{\partial \theta} < 0$ and the tax adjustment will further decrease the welfare beyond the direct effect ($\frac{\partial V}{\partial \theta}$). By the same token, a decrease in inequality will make for a welfare gain in the case of dominating necessities, and for a welfare-efficiency tradeoff to occur in the case of a dominance of luxuries. In general, higher inequality decreases aggregate welfare, whenever the social marginal utility of income is decreasing:[24] and this being the case, the possibility of

[23] Atkinson and Stiglitz (1980, p. 432) mention the possibility of a social marginal utility increasing in income.

[24] Formally, this follows from writing out (13)

$$\frac{dW}{d\theta} = \frac{d}{d\theta} \int_{\mathcal{Y}} \left[\lambda \tau(t, y) + v\left(p, y\right)\right] f(y, \theta) dy$$

where we can treat t and λ as constant by the envelope theorem. By the standard properties of second order stochastic dominance, $\frac{dW}{d\theta} < 0$ if $\lambda \tau(t, y) + v\left(p, y\right)$ is a concave function of y, which follows from decreasing γ, as $\frac{\partial^2}{\partial y^2}\left[\lambda \tau(t, y) + v\left(p, y\right)\right] = \frac{\partial \gamma}{\partial y}$.

the optimal tax reaction to higher inequality leading to higher welfare rests with the tax system being so progressive as to make social marginal utility increase.

3.3 Inequality and implicit taxation

So far we have considered optimal (indirect) tax adjustment in the face of an exogenous change in income distribution, and have discussed some implications of such an adjustment in terms of welfare. However, there are arguably many cases where governments are unable or unwilling to counter changing inequality by modifying indirect taxation. While in general with no tax adjustment the government's revenue constraint will be violated, a change in tax revenue will also occur, as the inequality shift will affect differently the demand functions for necessities and luxuries. This in turn implies that an inequality change will modify by itself the allocation of the tax burden across commodities – so that a change in inequality amounts to a (non optimal) implicit indirect tax.

To look at this, we may recall equation (6)

$$\frac{dQ_i^h}{Q_i} = \overline{\sigma}(\gamma, q_i/Q_i)$$

so that the relative change in compensated demand brought about by an unadjusted-for inequality change is given by

$$\frac{\partial}{\partial\theta}\frac{dQ_i^h}{Q_i}_{|dt=0,dA=0} = \frac{\partial}{\partial\theta}\overline{\sigma}(\gamma, q_i/Q_i) \tag{14}$$

This equation becomes more transparent if one assumes γ to be approximated linearly, in which case (14) reduces to

$$\frac{\partial}{\partial\theta}\frac{dQ_i^h}{Q_i}_{|dt=0,dA=0} = \frac{\alpha}{\lambda}\frac{\partial\mu_i}{\partial\theta} \tag{15}$$

where $\mu_i = \int_{\mathcal{Y}} y\varphi_i(y, p; \theta)dy$ is average income weighted by consumption shares and $\alpha = \partial\gamma/\partial y$. If the social marginal utility is decreasing, an income shock which raises μ_i brings about a decrease in the compensated demand for commodity i, and is in this sense akin to higher taxation for that commodity. This raises the question of the circumstances under which such is indeed the case. Sufficient conditions can be had from the following Proposition:

Proposition 5 *Assume* $\frac{\partial^2\gamma}{\partial y^2} = 0$ *and let an increase in* θ *signal an increase in inequality for given mean. Then* $\mathtt{sign}\left\{\frac{\partial}{\partial\theta}\left(dQ_i^h\right)\right\} = \mathtt{sign}\left\{\frac{\partial\gamma}{\partial y}\right\}$, *provided that (a)* q_i *is convex in income and such that* $\frac{\partial^2 q_i}{\partial y^2}/\frac{\partial q_i}{\partial y} < \frac{2}{\mu - y_{\min}}$, *or (b)* q_i *is concave in income and such that* $\frac{\partial^2 q_i}{\partial y^2}/\frac{\partial q_i}{\partial y} > -\frac{2}{y_{\max} - \mu}$.

Proof. See Appendix C ■

The main implication of Proposition 4 is that, under a linear approximation, if the social marginal utility of income is decreasing in income, an increase in inequality not countered by a tax adjustment amounts to raising taxation on a given commodity, whenever the corresponding Engel curve is not too convex or too concave. It should be noticed that with linear γ, λ will depend on mean income: this shows up in the relevant bounds of the above Proposition being wider, the lower the mean.

4 Concluding remarks

The distributional effects of indirect taxation are becoming an important topic for policy discussion, as governments are urged to shift taxation away from direct taxation. In this paper we provide a general framework to study the relationship between income distribution and optimal indirect taxation.

We assess how income distribution shapes the many-person Ramsey tax rule and Diamond's covariance, which we find to be always negative for normal goods when excise taxation is coupled with a poll income tax. We also express Diamond's covariance under a linear approximation which in principle may be useful for empirical investigations. Characterizing the solution to social welfare maximization allows us to assess the optimal reaction of commodity taxation to changes in income distribution, which turns out to depend on the progressivity of the tax system. This also leads to formalizing the conditions under which, following a marginal increase in inequality, a potential conflict arises between welfare and efficiency: the latter is the case if the system is regressive. If the system is progressive the optimal tax adjustments require an average decrease of the tax rates, combining social welfare gains with more equity. Finally, we conclude by pointing out how an increase in inequality not countered by an optimal tax adjustment amounts to an implicit taxation, as the implied change in the distribution of the tax burden across markets is driven by changes in the demand quantities.

We believe that this analysis may provide a framework which can be further developed in future research. Under this respect, a major suggestion is that the optimal taxation cost of internalising externalities, or controlling "internalities", can be different, according to the prevailing income distribution. Finally, this setting may be potentially fit for studying the links between direct and indirect taxation with different market power scenarios.

References

[1] Atkinson, A.B. (1970). On the measurement of inequality, Journal of Economic Theory, 2 (3), 244–263.

[2] Atkinson, A.B., and J.E. Stiglitz (1976) .The design of tax structure: direct vs. indirect taxation, Journal of Public Economics 6, 55-75

[3] Benassi, C. and A. Chirco (2006), Income share elasticity and stochastic dominance, Social Choice and Welfare, 26, 511 - 526.

[4] Boadway, R. (2012). From Optimal Tax Theory to Tax Policy: Retrospective and Prospective Views, 2009 Munich Lectures in Economics (Cambridge: MIT Press).

[5] Boadway, R. and P. Pestieau (2003). Indirect Taxation and Redistribution: The Scope of the Atkinson-Stiglitz Theorem, in Richard Arnott, Bruce Greenwald, Ravi Kanbur and Barry Nalebuff (eds.), Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz, (Cambridge, MA: MIT Press), 387–403.

[6] Browning, M. and C. Meghir (1991). The Effects of Male and Female Labor Supply on Commodity Demands, Econometrica, 59, 925-951.

[7] Castanheira, M., Nicodème G., Profeta P. (2012), On the political economics of tax reforms: survey and empirical assessment, International Tax and Public Finance 19, 598–624.

[8] Chetty, R. (2009a). Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods, Annual Review of Economics, 1, 451-488.

[9] Chetty, R. (2009b). Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance, American Economic Journal: Economic Policy, 1(2), 31-52.

[10] Diamond, P.A (1975). A Many-Person Ramsey Tax Rule, Journal of Public Economics, 4, 335-342.

[11] Diamond, P.A. and J.A. Mirrlees (1971). Optimal taxation and public production, American Economic Review 61, 8-27 and 261-78.

[12] Esteban J. (1986). Income share elasticity and the size distribution of income. International Econonomic Review 27, 439-444.

[13] European Commission (2014), Taxation trends in the European Union, Eurostat Statistical Books, http://ec.europa.eu/taxation_customs/resources/documents/taxation/gen_info/economic_analysis/tax_structures/2014/report.pdf

[14] European Commission (2016): Study and Reports on the VAT Gap in the EU-28 Member States:2016 Final Report-TAXUD/2015/CC/13. https://ec.europa.eu/taxation_customs/sites/taxation/files/2016-09_vat-gap-report_final.pdf.

[15] Feldstein, M. (1999). Tax Avoidance and the Deadweight Loss of the Income Tax, Review of Economics and Statistics, 81(4), 674-680.

[16] Kaplow, L. (2006). On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal, Journal of Public Economics, 90(6-7), 1235-50.

[17] Konishi, H., (1995). A Pareto-improving commodity tax reform under a smooth nonlinear income tax, Journal of Public Economics, 56, 413–446.

[18] Laroque, G.R. (2005). Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof, Economics Letters, 87(1), 141-4.

[19] Lambert, P. J. (1993). The Distribution and Redistribution of Income: A Mathematical Analysis, 2nd Edition. Manchester: Manchester University Press.

[20] Layard, R., G. Mayraz and S. Nickell (2008). The Marginal Utility of Income, Journal of Public Economics 92, 1846-1857.

[21] Myles, G.D. (2009a), Economic Growth and the Role of Taxation – Theory, OECD Economics Department Working Papers, 713.

[22] Myles, G.D. (2009b), Economic Growth and the Role of Taxation – Aggregate Data, OECD Economics Department Working Papers, 714

[23] Myles, G.D. (2009c), Economic Growth and the Role of Taxation – Disaggregate Data, OECD Economics Department Working Papers, 715.

[24] OECD (2014), Revenue Statistics 2014, OECD Publishing, http://dx.doi.org/10.1787/rev_stats-2014-en-fr.

[25] Piketty,T, Saez E. (2013). Optimal Labor Income Taxation, Handbook of Public Economics, Edited by Alan J. Auerbach, Raj Chetty, Martin Feldstein and Emmanuel Saez, Volume 5, 391-474.

[26] Piketty,T, Saez E. (2014,). Inequality in the long run, Science, 344, 838-843.

[27] Pirttilä, J. and I. Suoniemi (2014), Public Provision, Commodity Demand, and Hours of Work: An Empirical Analysis, The Scandinavian Journal of Economics 116, 1044–1067.

[28] Ramsey, F.P. (1927). A Contribution to the Theory of Taxation,The Economic Journal, 145, 47-61.

[29] Saez, E. (2002). The Desirability of Commodity Taxation under Non-linear Income Taxation and Heterogeneous Tastes, Journal of Public Economics, 83(2), 217-230.

[30] Saez, E., J. Slemrod, and S. Giertz (2012). The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review, Journal of Economic Literature 50(1), 3-50.

[31] Salanie, B. (2003), The Economics of Taxation, MIT press.

Appendix

A. Optimal taxation and the covariance matrix

Maximising the social welfare function (1) subject to the revenue constraint (2), optimal taxation (t, A) is identified by the twin first order conditions

$$\sum_j t_j \tfrac{\partial Q_j}{\partial p_i} + Q_i = \tfrac{1}{\lambda} \int_{\mathcal{Y}} \tfrac{\partial v}{\partial y} q_i(p, y_A) f(y, \theta) dy, \quad i = 1, ..., n \tag{A.1a}$$

$$1 + \sum_i t_i \frac{\partial Q_i}{\partial A} = \tfrac{1}{\lambda} \int_{\mathcal{Y}} \tfrac{\partial v}{\partial y} f(y, \theta) dy \tag{A.1b}$$

where we have used Roy's identity $\frac{\partial v}{\partial p_i} = -q_i \frac{\partial v}{\partial y}$, and λ is the associated multiplier. Equation (4) can be derived by noting that $\partial Q_i / \partial A = -\int_{\mathcal{Y}} \frac{\partial q_i}{\partial y} f(y, \theta) dy$, so that (A1.b) reduces to

$$1 - \int_{\mathcal{Y}} \tfrac{\partial \tau}{\partial y} f(y, \theta) dy = \tfrac{1}{\lambda} \int_{\mathcal{Y}} \tfrac{\partial v}{\partial y} f(y, \theta) dy$$

from which (4) follows by applying definition (3). As to equation (5), note that

$$\sigma(\gamma, q_i) = \int_{\mathcal{Y}} [\gamma(y, t) - \lambda] [q_i(p, y_A) - Q_i] f(y, \theta) dy \tag{A.2}$$

which can be cast in percentage terms as per (11)

$$\begin{aligned}\sigma(\gamma, q_i/Q_i) &= \int_{\mathcal{Y}} [\gamma(y,t) - \lambda] \left[\frac{q_i(p,y_A)}{Q_i} - 1\right] f(y,\theta)dy = \\ &= \int_{\mathcal{Y}} \gamma(t,y) [\varphi_i(y,p;\theta) - f(y;\theta)]\, dy\end{aligned}$$

where $\varphi_i(y,p;\theta) = q_i(p,y_A)f(y,\theta)/Q_i$ and $\int_{\mathcal{Y}} [\gamma(t,y) - \lambda] f(y,\theta)dy = 0$ by (3). Finally, by using Slutsky's equation to substitute for $\partial q_j/\partial p_i$, (A.1a) for commodity i boils down to[25]

$$\begin{aligned}\sum_j t_j \int_{\mathcal{Y}} \tfrac{\partial q_j^h}{\partial p_i} f(y,\theta)dy &= dQ_i^h = -Q_i + \int_{\mathcal{Y}} \tfrac{\partial \tau}{\partial y} q_i(p,y_A)f(y,\theta)dy \\ &\quad + \tfrac{1}{\lambda} \int_{\mathcal{Y}} \tfrac{\partial v}{\partial y} q_i(p,y_A)f(y,\theta)dy \\ &= -Q_i + \tfrac{1}{\lambda} \int_{\mathcal{Y}} \gamma(t,y_A)\, q_i(p,y_A)f(y,\theta)dy \qquad \text{(A.3)}\end{aligned}$$

Since by definition $\sigma(\gamma, q_i) = Q_i \sigma(\gamma, q_i/Q_i)$ and

$$\int_{\mathcal{Y}} \gamma(t,y_A)\, q_i(p,y_A)f(y,\theta)dy = \sigma(\gamma, q_i) + \lambda Q_i$$

one has

$$dQ_i^h = -Q_i + \tfrac{1}{\lambda} [Q_i \sigma(\gamma, q_i/Q_i) + \lambda Q_i]$$

from which equation (6) follows.

B. Proof of Proposition 1

(a) Integrating (5) by parts we get

$$\int_{\mathcal{Y}} \gamma(t,y) [\varphi_i(y,p;\theta) - f(y,\theta)]\, dy = -\int_{\mathcal{Y}} [\Phi_i(y,p;\theta) - F(y,\theta)] \frac{\partial \gamma}{\partial y} dy \quad \text{(A.4)}$$

where $\Phi_i(y,p;\theta) = \int_{y_{\min}}^{y} \varphi_i(x,p;\theta)\, dx$. Define now the Esteban (1986) elasticity of a density $f(y,\theta)$ as $\pi^f(y,\theta) = 1 + y\frac{\partial f}{\partial y}/f(y,\theta)$: we know from Benassi

[25] We substitute for $\frac{\partial q_j}{\partial p_i} = \frac{\partial q_j^h}{\partial p_i} - \frac{\partial q_j}{\partial y} q_i$, where $q_j^h = q_j^h(p, v(p,y))$. Using the symmetry of the Slutsky matrix, $dQ_i^h = \sum_j t_j \int_{\mathcal{Y}} \frac{\partial q_j^h}{\partial p_i} f(y,\theta)dy$ is the change in the compensated aggregate demand for good i.

and Chirco (2006) that, given any two differentiable densities f_i, $i=1,2$ with Esteban elasticity π^i, $\pi^1 > \pi^2$ for all $y \in \mathcal{Y}$ implies that f_1 first order stochastically dominates f_2, i.e. $F_1(y,\theta) \leq F_2(y,\theta)$ for all $y \in \mathcal{Y}$. It is easily seen that in our case

$$\pi^{\varphi_i}(y,p;\theta) = \pi^f(y,\theta) + \varepsilon_i$$

where $\varepsilon_i = \varepsilon_i(y,p)$ is the income elasticity of the individual demand function q_i. If commodity i is normal $\pi^{\varphi_i} - \pi^f > 0$ for all $y \in \mathcal{Y}$, so that $\Phi_i(y,p;\theta) - F(y,\theta) \leq 0$ (reverse inequalities obviously apply for inferior commodities), from which the result follows.

(b) From(5), (6), and (A.4) one has

$$\begin{aligned} \frac{dQ_i^h}{Q_i} - \frac{dQ_j^h}{Q_j} &= \tfrac{1}{\lambda}\int_{\mathcal{Y}} \gamma(t,y)\left[\varphi_i(y,p;\theta) - \varphi_j(y,p;\theta)\right] dy \\ &= -\tfrac{1}{\lambda}\int_{\mathcal{Y}} \left[\Phi_i(y,p;\theta) - \Phi_j(y,p;\theta)\right] \frac{\partial \gamma}{\partial y} dy \end{aligned}$$

From **(a)** we know that $\pi^{\varphi_i} - \pi^{\varphi_j} > 0$ for all $y \in \mathcal{Y}$ implies $\Phi_i(y,p;\theta) - \Phi_j(y,p;\theta) \leq 0$, and that $\pi^{\varphi_i} - \pi^{\varphi_j} = \varepsilon_i(y,p) - \varepsilon_j(y,p)$; since $\frac{\partial\gamma}{\partial y} < 0$ the result follows.

C. Proof of Proposition 4

From (A.4), $\lambda dQ_i^h = \sigma(\partial v/\partial y, q_i) = -\lambda Q_i + \int_{\mathcal{Y}} \gamma(t,y_A)\, q_i(p,y_A) f(y,\theta) dy$. With $\lambda > 0$, the standard properties of second order stochastic dominance for given mean imply that

$$\mathtt{sign}\left\{\frac{\partial}{\partial\theta}\left(dQ_i^h\right)\right\} = \mathtt{sign}\left\{\left[\gamma(t,y_A) - \lambda\right]\frac{\partial^2 q_i}{\partial y^2} + 2\frac{\partial q_i}{\partial y}\frac{\partial \gamma}{\partial y}\right\}$$

To ease notation, let $y_{A\min} = y_{\min} - A > 0$, $y_{A\max} = y_{\max} - A$, and $\mu_A = \mu - A$. Consider now case (a). Notice that $\mathtt{sign}\left\{\frac{\partial\gamma}{\partial y}\right\} = \mathtt{sign}\left\{\gamma(t,y_A) - \gamma(t,y_{A\min})\right\}$, and, since λ is the expected value of γ, $\mathtt{sign}\left\{\frac{\partial\gamma}{\partial y}\right\} = \mathtt{sign}\left\{\lambda - \gamma(t,y_{A\min})\right\}$, so that

$$2\frac{\partial \gamma}{\partial y} / \left[\lambda - \gamma\left(t, y_{A\,\mathrm{min}}\right)\right] > 0$$

Let now $B = \left[\gamma\left(t, y_A\right) - \lambda\right] \frac{\partial^2 q_i}{\partial y^2} + 2 \frac{\partial q_i}{\partial y} \frac{\partial \gamma}{\partial y}$ and $B_{\mathrm{min}} = \left[\gamma\left(t, y_{A\,\mathrm{min}}\right) - \lambda\right] \frac{\partial^2 q_i}{\partial y^2} + 2 \frac{\partial q_i}{\partial y} \frac{\partial \gamma}{\partial y}$: then either $\frac{\partial \gamma}{\partial y} < 0$ and $B < B_{\mathrm{min}}$ or $\frac{\partial \gamma}{\partial y} > 0$ and $B > B_{\mathrm{min}}$. The condition

$$\frac{\partial^2 q_i}{\partial y^2} / \frac{\partial q_i}{\partial y} < 2\frac{\partial \gamma}{\partial y} \left[\lambda - \gamma\left(t, y_{A\,\mathrm{min}}\right)\right]$$

ensures $B_{\mathrm{min}} < 0$ in the former case and $B_{\mathrm{min}} > 0$ in the latter, and hence it provides an upper bound on (relative) convexity such that $\mathtt{sign}\left\{\frac{\partial \gamma}{\partial y}\right\} = \mathtt{sign}\left\{\frac{\partial}{\partial \theta}\left(dQ_i^h\right)\right\}$. If γ is linear in income, $\frac{\partial \gamma}{\partial y} = \alpha\left(t\right)$ is a constant such that $\gamma\left(t, y_A\right) = \alpha_0(t) + \alpha\left(t\right) y_A$; hence, using (4)

$$\gamma\left(t, y_{A\,\mathrm{min}}\right) - \lambda = \left(y_{\mathrm{min}} - \mu\right)\alpha$$

so that

$$2\frac{\partial \gamma}{\partial y} / \left[\lambda - \gamma\left(t, y_{A\,\mathrm{min}}\right)\right] = \frac{2}{\mu - y_{\mathrm{min}}}$$

Similarly for case (b), $\mathtt{sign}\left\{\frac{\partial \gamma}{\partial y}\right\} = \mathtt{sign}\left\{\gamma\left(t, y_{A\,\mathrm{max}}\right) - \gamma\left(t, y_A\right)\right\}$, and, since λ is the expected value of γ, $\mathtt{sign}\left\{\frac{\partial \gamma}{\partial y}\right\} = \mathtt{sign}\left\{\gamma\left(t, y_{A\,\mathrm{max}}\right) - \lambda\right\}$, so that

$$2\frac{\partial \gamma}{\partial y} / \left[\lambda - \gamma\left(t, y_{A\,\mathrm{max}}\right)\right] < 0$$

Let $B_{\mathrm{max}} = \left[\gamma\left(t, y_{A\,\mathrm{max}}\right) - \lambda\right] \frac{\partial^2 q_i}{\partial y^2} + 2 \frac{\partial q_i}{\partial y} \frac{\partial \gamma}{\partial y}$: then either $\frac{\partial \gamma}{\partial y} < 0$ and $B < B_{\mathrm{max}}$ or $\frac{\partial \gamma}{\partial y} > 0$ and $B > B_{\mathrm{max}}$. The condition

$$\frac{\partial^2 q_i}{\partial y^2} / \frac{\partial q_i}{\partial y} > 2\frac{\partial \gamma}{\partial y} / \left[\lambda - \gamma\left(t, y_{A\,\mathrm{max}}\right)\right]$$

ensures $B_{\mathrm{max}} < 0$ in the former case and $B_{\mathrm{max}} > 0$ in the latter, and hence it provides a lower bound on (relative) concavity such that $\mathtt{sign}\left\{\frac{\partial \gamma}{\partial y}\right\} = \mathtt{sign}\left\{\frac{\partial}{\partial \theta}\left(dQ_i^h\right)\right\}$. Again, one can write

$$2\frac{\partial \gamma}{\partial y} / \left[\lambda - \gamma\left(t, y_{A\,\mathrm{max}}\right)\right] = -2 / \left(y_{\mathrm{max}} - \mu\right)$$

by using the fact that under linearity $\gamma\left(t, y_A\right)=\alpha_0(t)+\alpha\left(t\right) y_A$.