Optimal Policy with Heterogeneous Preferences

Louis Kaplow*

Abstract

Optimal policy rules – including those regarding income taxation, commodity taxation, public goods, and externalities – are typically derived in models with preferences that are homogeneous. This article reconsiders many central results for the case in which preferences for commodities, public goods, and externalities are heterogeneous. When preference differences are observable, standard second-best results in basic settings are unaffected, except those for the optimal income tax. Optimal marginal income tax rates may be higher or lower on types who derive more utility from various goods, depending on the nature of preference differences and the concavity of the social welfare function. When preference differences are unobservable, all policy rules may change. The determinants of even the direction of optimal rule adjustments are many and subtle.

JEL Classes D61, D62, D63, H21, H23, H24, H43, K34

Keywords: heterogenous preferences, optimal income taxation, commodity taxation, externalities, public goods, social welfare

*Harvard University and National Bureau of Economic Research.

1. Introduction

Most analytical work in public economics that derives rules for optimal policy assumes that preferences are homogeneous. This characterization applies, for example, to most of the literature on optimal income taxation, commodity taxation, public goods provision, and the control of externalities. The reason for this simplification is tractability; indeed, the second-best problems are complex even when preferences are stipulated to be homogeneous.

It is natural to explore whether and how standard results are modified when preferences are heterogeneous. There may be ordinary differences in tastes: Some individuals may prefer chocolate, others vanilla; some may love nature and thus highly value improvements to national parks or pollution regulations that restore wilderness habitats, whereas others may desire city life and accordingly prefer enhancements to urban amenities. More particular differences may arise on account of physical and mental disabilities.

These possibilities raise a number of questions. Does Atkinson and Stiglitz's (1976) result on uniform commodity taxation, the Samuelson rule for public goods, or the Pigouvian dictum that corrective taxes and subsidies should equal the marginal harms and benefits (respectively) from externalities survive the introduction of heterogeneity, and if not how do the results change? Likewise, how are standard first-order conditions for optimal income taxation affected, including in the case in which some preference differences may be observable? Finally, how do the answers to such questions depend on the manner in which preferences differ? (In other words, in what respects is the problem of preference heterogeneity itself a heterogeneous phenomenon?)

A related, more focused set of questions may be directed at a particular set of results in the literature. Certain policy rules have been demonstrated to be optimal because following them generates Pareto improvements, one consequence of which is that the optimal rules are independent of the particular choice of social welfare function (SWF). This approach is featured in generalizations of Atkinson and Stiglitz (1976) on commodity taxation as well as in work on public goods and on externalities. See, for example, Kaplow (1996, 2006a, 2006b), Konishi (1995), and Laroque (2005). Yet when preferences are heterogeneous, Pareto improvements are highly unlikely. Virtually any policy reform, no matter how desirable, is likely to generate some losers because certain individuals will have idiosyncratic preferences that are better served by the inefficient status quo. For example, even substantial reductions of a highly damaging pollutant at minimal cost may disfavor some whose bodies are insensitive to the pollutant and who especially enjoy polluting activities that need to be curtailed. Ng (1984) has suggested that, as a consequence of heterogeneity, we might instead ask whether policies offer a "quasi-Pareto" improvement, by which he means a gain, on average, to individuals at every level of income (thereby addressing standard distributive concerns). Nevertheless, it is appropriate to explore formally and systematically how, if at all, the derivations and results in this literature need to be adjusted in light of preference heterogeneity.

This article undertakes a preliminary exploration of these questions. Section 2 presents a model with a nonlinear income tax and commodity taxes where individuals' preferences for commodities are heterogeneous. In section 3, the model is analyzed for the case in which preference differences are observable. This case is of interest because it is analytically tractable, facilitates examination of the case with unobservable heterogeneity, and has some elements of realism (for example, certain physical disabilities, with associated differences in preferences, are observable). It turns out that the Atkinson-Stiglitz result concerning the optimality of uniform commodity taxation and a variety of extensions are preserved. By contrast, results on optimal income taxation differ. For individuals with preferences that yield higher utility for a given level of disposable income, optimal marginal tax rates may be higher or lower, depending on the nature of preference differences and on the concavity of utility functions and of the SWF.

Section 4 sketches a number of extensions. When preference differences are unobservable, uniform commodity taxation is no longer desirable. However, determinants of the direction and magnitude of deviations are complex and subtle. The basic explanation for the results is that differentials in commodity taxation are optimal because (and to the extent that) they have effects similar to those of the adjustments to the income tax (characterized in section 3) that would have been optimal if preference differences were observable. This section also extends the results to policy rules for public goods and externalities, for which many of the conclusions are analogous to those for commodity taxation: When preferences differences are observable, benchmark results on first-best policy rules (the Samuelson rule for public goods and the Pigouvian prescription for complete internalization of externalities) continue to hold. When preference differences are not observable, deviations depend on some (but not all) of the same sorts of factors as with commodity taxation and serve the same purpose of indirectly substituting for the redistributive adjustments to the income tax that would have been optimal in the case of observability. Concluding remarks are offered in section 5.

Prior work on heterogeneity is of a number of types. Saez (2002) introduces heterogeneity regarding previously established deviations from the Atkinson-Stiglitz uniform commodity tax result: Atkinson and Stiglitz (1976) had shown that relaxing labor separability favors nonuniformity and Mirrlees (1976) had shown that relaxing the assumption that preferences are independent of earning ability favors nonuniformity; Saez (2002) demonstrates that when these relationships are not deterministic but rather reflect mere correlations, similar results obtain.[1] By contrast, the present article abstracts from these two considerations and examines instead effects of heterogeneity that pertain more directly to the marginal social value of redistribution than to the labor-leisure distortion caused by income taxation.

Some discussion of these other sorts of heterogeneity appears in prior literature – especially on public goods, e.g., Hylland and Zeckhauser (1979), Ng (1984), Boadway and Keen (1993), and Kaplow (1996) – but does not suggest or demonstrate the results derived here. Other papers have examined certain technical aspects of the nonlinear income tax problem when additional dimensions are introduced. See Ebert (1988) and Tarkiainen and Tuomala (1999). Tarkiainen and Tuomala (1999) also present some simulations, but their example has special features and the means of the parameters in their two-dimensional case differ (often significantly) from the values in the one-dimensional case, so it is hard to interpret the simulation results regarding differences in optimal redistribution in the presence of heterogeneity. Other

[1]Saez (2002) notes the existence of other forms of preference heterogeneity but does not focus on them.

work, notably Boadway et al. (2002), focuses on which self-selection constraints are binding in a model in which preferences are of two types.

Finally, some additional papers introduce specific forms of heterogeneity in certain variations of the income or commodity tax problem. Blackorby and Donaldson (1988) find that in-kind provision of medical care – or, if not feasible, differential taxation of medical care and other goods – may be optimal when there are unobservable differences in medical needs. Sandmo (1993) explores some optimal income tax implications of heterogeneous tastes for work in a model in which there are no differences in earning abilities and only one type of consumption. Marchand, Pestieau, and Racionero (2003) consider differences in the source of disutility to labor that are unobservable, are relevant to social welfare (a form of nonwelfarist SWF), and are related differently to demands for particular commodities. Hellwig (2004) addresses optimal pricing by a public monopolist with heterogenous consumers. And Fleurbaey (2006) introduces preference heterogeneity along with a nonwelfarist SWF that features a laissez-faire criterion that favors noninterference with consumption choices.

As a whole, however, prior work does not address most of the questions examined here or how the answers depend on observability or on the policy context (commodity taxation, public goods, externalities, or purely income taxation). Of particular note, past inquiries (including by this author) typically speak of heterogeneity as if it were a unitary phenomenon when making general conjectures about how results may change with heterogeneity or when examining particular models. By contrast, the present analysis allows preferences among commodities (or public goods or external effects) to differ in various ways and finds that whether results change and, when they do, in what direction depends on the nature of the heterogeneity that is introduced.

2. Model

There are n commodities, indexed by i; a particular commodity is denoted x_i. Individuals choose labor effort, l, and commodity vectors, x, to maximize utility $u(x, l, \theta)$, where θ is a vector of preference parameters. The consumer price vector is q, taken to be the sum of fixed producer prices p and a commodity tax vector τ (commodity taxes may be negative, but q must be positive). The nonlinear income tax $T(wl)$ is a function of individuals' income, wl, the product of their unobservable wage rate (ability level) w and labor effort. The budget constraint is

$$(2.1)\ qx = wl - T(wl).$$

The government chooses the commodity tax vector τ (equivalently, the consumer price vector q) and the nonlinear income tax T to maximize social welfare, given by

$$(2.2)\quad SW = \int W(v(w,\theta,\tau,T))dF(w,\theta),$$

where W is a concave function, v is an individual's indirect utility function – indicating the maximized value of u for ability level w and preference parameters θ when the tax regime (τ, T) is taken as given – and F is the distribution function. There is a revenue constraint – commodity tax plus income tax revenue must meet a fixed target R – that does not need to be examined explicitly here.

Much of the analysis will focus on the case in which the preference parameters θ are observable. Accordingly, the pertinent nonlinear income tax schedules are $T(wl, \theta)$; that is, for each preference type θ, there is a separate income tax schedule. It will be assumed that commodity taxes, τ, cannot be type-specific because of the possibility of resale. This restriction often will not matter; comments in footnotes will discuss implications when it does.

Taking commodity taxes as given, the optimal nonlinear income tax is characterized by a first-order condition derived, for example, in Atkinson and Stiglitz (1980), who make various simplifications (notably, that utility is separable between consumption and labor effort) and offer other caveats that will not be examined further here. Their condition can be expressed as:

$$(2.3)\quad \frac{T'(wl,\theta)}{1-T'(wl,\theta)} = \frac{\int_w^\infty \frac{v_c(w,\theta)}{v_c(\omega,\theta)}\left(1-\frac{W'(v(\omega,\theta))v_c(\omega,\theta)}{\lambda}\right)f(\omega,\theta)d\omega}{\varepsilon w f(w,\theta)}.$$

Primes denote derivatives, v_c is the marginal utility of consumption, f is the density function derived from F, λ is the marginal social value of a dollar (i.e., the shadow price on the government's revenue constraint), and ε (implicitly a function of w and θ) is related to the labor supply elasticity. (In stating expression (2.3), the fact that indirect utility, v, depends on τ, which is taken as given, and T is suppressed.)

The left side of expression (2.3) indicates the marginal income tax rate as a fraction of the untaxed proportion of marginal earnings. On the right side, the denominator indicates the marginal distortionary cost of a higher marginal income tax rate. It is the product of three components: ε (as noted, a term related to the labor supply elasticity); w, which measures the productivity lost by each unit reduction in labor supply (and likewise indicates the marginal revenue loss); and $f(w, \theta)$, which is the portion of the population that is distorted at the margin.

The integral in the numerator is related to the benefit from higher marginal income tax rates in terms of revenue raised on those earning more than wl, for whom a marginal income tax rate increase at income wl is inframarginal. (This integral is from the marginal type, w, who faces the marginal rate increase, to the upper limit of the distribution of ability types, taken here to be unbounded; ω is the variable of integration, indexing types with ability above w.) The key term for present purposes is that in the large parentheses in the integrand. The numerator of the latter component, $W'v_c$, is the marginal contribution to social welfare caused by a unit increase in utility of the pertinent type multiplied by that type's marginal utility of consumption – i.e., net-of-income-tax income or disposable income. This term is divided by λ, which converts units of social welfare into dollars. Accordingly, as the overall marginal contribution of individuals' consumption to social welfare is greater (ceteris paribus), the term in parentheses will be smaller, the value of the integral will be lower, and thus the optimal marginal tax rate will be lower.

It is important to keep in mind that, for the case in which θ is observable, this condition is separately stated for each θ. These separate conditions are linked to the single social optimization problem by the common shadow price λ. When comparing groups with different values of θ, it will be true, ceteris paribus, that those with higher marginal social valuations of consumption should be subject to lower marginal income tax rates. This "ceteris paribus" statement is, however, highly problematic because θ will affect other components in expression (2.3) as well. (Notably, a given ability type w with a different θ may choose to supply a different level of labor effort, l, and thus earn a different level of income, wl.) Nevertheless, the standard practice of interpreting this sort of first-order condition term by term provides some valuable insight into this highly complicated problem.

The analysis will focus on utility functions that are at least weakly separable in labor. Specifically, attention will be confined to forms of $u(x, l, \theta)$ that can be written as $u(u_1(x, \theta), l, \theta)$. This restriction is meaningful because x is a vector. This weak separability assumption means that, in allocating resources among the x_i's, it is immaterial how much labor effort, l, was required to produce the individual's level of disposable income. Likewise, the allocation among the x_i's will not directly influence the choice of labor effort, l; all that will matter is the level of subutility, $u_1(x, \theta)$, thereby obtained.

To examine more specifically the effects of particular sorts of preference heterogeneity on the optimal income and commodity tax problem, it often is useful to explore a more specific utility function. The functional form to be considered will be a five-parameter (by n) function, $\theta = (\alpha, \beta, \gamma, \delta, \rho)$, where each of these parameters may take on a different value for each commodity. This utility function is:

$$(2.4)\ u(x,l,\alpha,\beta,\gamma,\delta,\rho) = \sum_{i=1}^{n}\left(\alpha_i \frac{(\beta_i x_i + \gamma_i)^{1-\rho_i}}{1-\rho_i} + \delta_i\right) - z(l).$$

It is assumed that $\alpha_i > 0$, $\beta_i > 0$, $\rho_i > 0$ (each for all i), $z' > 0$, and $z'' > 0$. (For the case in which the concavity parameter, ρ_i, equals 1, the fractional expression is replaced by $\ln(\beta_i x_i + \gamma_i)$; it will be apparent below that using this alternative will yield the same first-order conditions, so none of the analysis is affected.) It is helpful for most discussion to think of a base case in which, for any commodity x_i, $\alpha_i = \beta_i = 1$ and $\gamma_i = \delta_i = 0$. Relative to that case, a higher α_i magnifies the contribution of x_i to utility without being subject to the curvature effect through ρ_i; a higher β_i magnifies the contribution of x_i to utility but is subject to the curvature effect through ρ_i; γ_i adds to utility in a manner that is subject to the curvature effect through ρ_i, and δ_i adds to utility without being subject to the curvature effect through ρ_i. It is obvious that raising any of these four parameters raises utility; however, as will be explored, they have a qualitatively different effect on marginal utility and thus may have a different effect on optimal income tax rates.

3. Analysis

3.1. Optimal Commodity Taxation

For the case in which individuals' preference parameters are observable, the existence of preference heterogeneity has no significant impact on results concerning optimal commodity taxation in the presence of a nonlinear income tax. Specifically, weak separability of labor is sufficient to generate the result that no differentiation is optimal, which implies that optimal commodity taxes may be taken to be zero. This result does not require that the income tax be optimal, and it can be extended to show the optimality of various partial commodity tax reforms, such as moving differentiated commodity taxes proportionally toward uniformity.

The proof of all of these results in Kaplow (2006a) (and in some other work, e.g., Laroque (2005)) requires only weak separability of labor and homogenous preferences (in particular, that the subutility function of commodities, u_1 in the above formulation, be common). However, when preference parameters are observable, the problem can be analyzed as if preferences are identical because the nonlinear income tax can be preference-type-specific.

Rather than reconstructing the pertinent proofs, it should be sufficient to review their two key steps. The first step – the one that depends on homogenous preferences – involves constructing an adjustment to the preexisting (arbitrary, i.e., not necessarily optimal) nonlinear income tax so that, when combined with the commodity tax reform (say, a move toward uniformity), everyone's utility is held constant. Specifically, this intermediate tax schedule, $T°(wl)$, is defined such that $V(\tau, T, wl) = V(\tau^*, T°, wl)$ for all wl – where V is an indirect subutility function indicating the maximized value of what is here denoted u_1. For this construction to be feasible – i.e., for the same tax schedule to preserve subutility for all individuals – it is necessary that the underlying subutility functions are identical (but homogeneity in other respects is not required).[2] Here, u_1 depends on θ. However, when θ is observable, a separate income tax schedule is applied for each θ; hence, a separate income tax adjustment may be employed as well. Thus, for each θ, one can define $T°(wl, \theta)$ as the adjustment to $T(wl, \theta)$ such that $V(\tau, T, wl, \theta) = V(\tau^*, T°, wl, \theta)$ for all wl.

At this point, the remainder of the proofs in Kaplow (2006a) and in other pertinent papers goes through. In particular, it was shown that this tax adjustment that holds utility constant for each level of labor supply will in fact induce individuals to choose the same level of labor supply. (The reasoning is, in essence, that the tax adjustment, when combined with the contemplated commodity tax adjustment, produces the same mapping from l to total utility as was produced initially, so whatever l maximized utility initially will continue to do so.) To

[2]For further discussion, consider Boadway and Keen's (1993) examination of preference heterogeneity in the public goods context (in an analysis that employs first-order conditions for social welfare maximization rather than constructing Pareto improvements).

complete the proofs, it was demonstrated that commodity tax reforms that were efficient in the narrow, conventional sense – including moves toward uniformity – produce a revenue surplus. (The intuition is that, since the income tax adjustment holds utility constant, yet there is an efficiency gain, it must be that the income tax adjustment is absorbing the dollar equivalent to that underlying efficiency gain.) This surplus, in turn, could be rebated pro rata, generating a Pareto improvement.

Upon reflection, the foregoing result is not surprising. Results in in the homogeneous case carry over in the presence of heterogeneity when preference differences are observable because one can decompose the latter problem into a number of cases of the former. Furthermore, because the proof technique employed in the literature in question does not depend on the optimality of the income tax (which is affected by heterogeneity) or any global properties of the system, the fact that these features may change in the presence of preference heterogeneity does not affect these results regarding optimal commodity taxation.[3]

3.2. Optimal Income Taxation

With observable differences in preferences, it is optimal to employ different nonlinear income tax schedules. For the present analysis, commodity taxes can be put to the side; specifically, they will be taken to be zero. The focus will be on how the marginal social contribution of disposable income to social welfare, $W'v_c$, depends on θ when preferences have the form given by expression (2.4).[4] As already noted, it is obvious that utility *levels* rise with α, β, γ, and δ. To confirm this, note that the partial derivative of u is clearly positive for each of these parameters, and the total derivative will be the same (the envelope theorem). If the SWF is utilitarian, which means that W' is constant, this feature is of no consequence. If instead the SWF is strictly concave, i.e., $W'' < 0$, then a group with a higher value of any of the parameters should, ceteris paribus, face higher marginal income tax rates because the marginal contribution of their utility to social welfare is lower on account of their already being better off.[5]

Next, consider how raising any of these four parameters affects the *marginal* utility of

---

[3]Note that the requirement that commodity taxes be anonymous (i.e., not type-specific) is not restrictive because all types should face the same (zero) tax vector. If one relaxed the labor separability assumption, however, the optimal commodity tax problem would complicated by the fact that the optimal differentiation, which reflects the degree to which various commodities are complements to or substitutes for labor, could depend on individuals' specific preferences. (Suppose, for example, that x_i was a complement and x_j was a substitute for type θ_1 but the opposite was true for type θ_2.) Ideally, each type would be subject to distinctive commodity taxes and subsidies, but if this is infeasible the optimum would (put crudely) reflect a weighted average of what would be optimal for different types considered separately. When the optimum for each type is the same, which is true with separability, this problem does not arise.

[4]Such analysis is incomplete because of other aspects of the social first-order condition for optimal marginal income tax rates (2.3) that may differ, other standard caveats in interpreting first-order conditions to make conjectures about the optimum, and difficulties of comparing optimal income tax schedules across groups because individuals of a given earning ability w may choose to exert different labor effort l and thus earn different income wl if their preferences differ.

[5]Higher (lower) optimal marginal income tax rates will also tend to be associated with lower (higher) optimal grants (i.e., y-intercepts of the T schedule, $-T(0, \theta)$, the value of the transfer to individuals in a given group who earn no income).

consumption, v_c, which will be relevant to the marginal social welfare gain from additional disposable income regardless of the concavity of W (short of maximin, wherein the welfare weight on everyone but the least-well-off individual is zero).

Effects on marginal utility are more varied. To analyze them, begin with individuals' first-order conditions. When expression (2.4) is maximized subject to the budget constraint, we have, for each x_i,

$$(3.1)\ \alpha_i \beta_i (\beta_i x_i + \gamma_i)^{-\rho_i} = \mu q_i,$$

where μ is the marginal utility of disposable income (the shadow price on the budget constraint; also equal to v_c, employing the notation of the indirect utility function from expressions (2.2) and (2.3)). It is useful to restate these conditions as

$$(3.2)\ x_i = \frac{\left(\frac{\alpha_i \beta_i}{\mu q_i}\right)^{1/\rho_i} - \gamma_i}{\beta_i}.$$

Expression (3.2) indicates how each of the x_i's is a function of the preference parameters for the corresponding good, the good's consumer price (which is taken here to be constant), and the marginal utility of consumption, μ. The first-order condition for labor effort is

$$(3.3)\ z'(l) = w\big(1 - T'(wl)\big)\mu.$$

The task is to determine how μ changes with each of the preference parameters. Let a generic preference parameter (aside from ρ) for good k be denoted ϕ_k. Differentiating the budget constraint (2.1) with respect to ϕ_k yields

$$(3.4)\ q_k \frac{\partial x_k}{\partial \phi_k} + \sum_{i=1}^{n} \left(q_i \frac{\partial x_i}{\partial \mu} \frac{d\mu}{d\phi_k} \right) = w(1 - T') \frac{dl}{d\mu} \frac{d\mu}{d\phi_k}.$$

Rearranging terms allows us to state

$$(3.5)\ \frac{d\mu}{d\phi_k} = \frac{q_k \frac{\partial x_k}{\partial \phi_k}}{w(1 - T') \frac{dl}{d\mu} - \sum_{i=1}^{n} \left(q_i \frac{\partial x_i}{\partial \mu} \right)}.$$

To determine the sign of expression (3.5), begin with the denominator. Differentiating the first-order condition for labor effort, l, expression (3.3), with respect to μ, and rearranging terms, produces

$$(3.6)\ \frac{dl}{d\mu} = \frac{w(1 - T')}{z'' + \mu w^2 T''} > 0.$$

The inequality in expression (3.6) follows because the denominator must be positive according to the second-order condition for the choice of l (the denominator is the negative of the second derivative of the individual's Lagrangian with respect to l). Differentiating expression (3.2) with respect to μ yields

$$(3.7)\ \frac{\partial x_i}{\partial \mu} = -\frac{\alpha_i}{\rho_i q_i \mu^2}\left(\frac{\alpha_i \beta_i}{\mu q_i}\right)^{\frac{1-\rho_i}{\rho_i}} < 0.$$

Expressions (3.6) and (3.7) together imply that the denominator of expression (3.5) is positive. Accordingly, the sign of $d\mu/d\phi_k$ is the same as the sign of $\partial x_k/\partial \phi_k$.

This result is in accord with intuition. Suppose, for example, that $\partial x_k/\partial \phi_k > 0$. In this case, the direct effect of raising ϕ_k is to induce the individual to purchase more of x_k, which through the budget constraint requires some combination of reductions in expenditures on the other x_i's ($i \neq k$) and an increase in l, both of which imply a higher μ, which is apparent from expressions (3.2) and (3.3). (The same logic holds, mutatis mutandis, if $\partial x_k/\partial \phi_k < 0$.)

To sign the $d\mu/d\phi_k$'s, therefore, all that remains is to sign the $\partial x_k/\partial \phi_k$'s, which is straightforward from expression (3.2).

$$(3.8)\ \frac{\partial x_i}{\partial \alpha_i} = \frac{\beta_i x_i + \gamma_i}{\rho_i \alpha_i \beta_i} > 0.$$

$$(3.9)\ \frac{\partial x_i}{\partial \beta_i} = \frac{x_i}{\beta_i}\frac{1 - \rho_i}{\rho_i} + \frac{\gamma_i}{\rho_i \beta_i^2}.$$

$$(3.10)\ \frac{\partial x_i}{\partial \gamma_i} = -\frac{1}{\beta_i} < 0.$$

$$(3.11)\ \frac{\partial x_i}{\partial \delta_i} = 0.$$

Expression (3.8) implies that $\partial\mu/\partial\alpha_i > 0$. Because $v_c = \mu$, it follows that the overall effect on the social marginal value of consumption, $W'v_c$, of raising α_i is ambiguous: utility rises, so W' falls if W is strictly concave; however, marginal utility rises, so v_c rises. Those with a higher α_i get more out of consuming x_i, which has these two competing effects.

Although one might have expected the effect of β_i to be qualitatively similar, this is not the case. Expression (3.9) indicates that the sign of $\partial\mu/\partial\beta_i$ is ambiguous. For convenience, interpretation will be confined to the benchmark case in which $\gamma_i = 0$. (When this is not so, the critical value of ρ_i for which the sign of the expression reverses would be adjusted accordingly.) If $\rho_i < 1$, then $\partial\mu/\partial\beta_i > 0$, and the results are indeed like those in the prior case. However, if $\rho_i > 1$, then $\partial\mu/\partial\beta_i < 0$, and the marginal utility effect combines with the effect on W' to reduce the social marginal valuation of consumption. The difference arises because the β_i coefficient directly multiplies x_i and thus is subject to the concavity of the subutility function, and when that concavity is sufficiently high ($\rho_i > 1$), the diminishing returns effect dominates the efficiency effect in determining how marginal utility changes.

Combining these two results, it can be seen that there are two senses in which an individual might be seen to get "more" out of a good x_i than do others. In the former case, the *contribution to utility* of good x_i is multiplied by the factor α_i. In the latter case, there is a sense in which the *effective quantity* of good x_i is multiplied by the factor β_i. One might interpret the former as an individual enjoying a good more and the latter as an individual being able to use a good more effectively. These notions are similar, but as just indicated they are not the same.

Expression (3.10) indicates that $\partial\mu/\partial\gamma_i < 0$. A higher γ_i is thus associated with higher utility and lower marginal utility, which unambiguously reduces the marginal social welfare weight $W'v_c$. A higher γ_i may be interpreted as it being as if the individual is naturally endowed with some of the good before purchasing any on the market. Or, put another way, a low γ_i – say, a negative value, relative to a benchmark value of 0 – would be a type of disability, wherein the individual needs to purchase some amount of x_i to reach the same starting point as others. Clearly, an individual with such a low γ_i will have both lower total utility and higher marginal utility, which unambiguously implies a higher social welfare weight.

Finally, as expression (3.11) indicates (and is obvious), $\partial\mu/\partial\delta_i = 0$. A higher δ_i implies a higher utility level but no difference in marginal utility, and thus decreases the social welfare weight $W'v_c$ if and only if the SWF is strictly concave. Indeed, as is apparent from expression (2.4), the parameter δ_i, despite its subscript, is not commodity-specific. All of the δ_i's might be aggregated into a single parameter δ for present purposes.

To summarize, the analysis in this subsection reveals that preference differences that all imply higher *utility levels* can nevertheless have qualitatively different implications for individuals' *marginal utility* of consumption. Moreover, marginal utility is directly relevant to the marginal social welfare weight and thus in determining how preference differences should affect optimal nonlinear income taxation. Indeed, with a utilitarian SWF, it is only the effect on marginal utility that matters. With strictly concave welfare functions, the utility level matters as well, and depending on the type of difference in preference, its strength, and the concavity of utility and of the SWF, individuals who have preferences that generate greater utility for a given level of disposable income may receive higher or lower marginal social welfare weights on their disposable income and thus optimally be subject to lower or higher marginal income tax rates.[6]

4. Extensions

4.1. Optimal Commodity Taxation with Unobservable Types

The analysis in section 3 takes preferences to be observable, whereas in this subsection preferences differences will be assumed to be entirely unobservable.[7] As a consequence, it will no longer be possible to redistribute across preference types through the income tax. When θ is not observed, the income tax schedule T in expression (2.3) must be the same for all types. The terms on the right side in this first-order condition will accordingly represent weighted averages of sorts. Notably, $W'v_c$ will be the product of the social welfare weight and marginal utility of consumption, averaged for all types at each pertinent level of earnings, wl.[8]

This limitation on across-preference-type redistribution through type-specific T schedules means that (even with weak labor separability) there is a potential role for differential commodity taxation to improve social welfare. Specifically, if individuals who have above-average demands for some commodity x_i would ideally (i.e., if types were observable) be subject to higher (lower) income taxation, then to some extent it will be optimal to tax (subsidize) that commodity relative to others . Favorable redistribution would result, and starting from the point of uniform commodity taxation there would be no first-order loss from consumption distortion. The optimal level of commodity taxes and subsidies would reflect a redistribution-distortion tradeoff, where here the redistribution is "horizontal" (across preference types) and the distortion is of commodity demands (rather than of the labor-leisure choice).

To explore which commodities should be taxed or subsidized, it is necessary first to ascertain the relationship between commodity demands and preferences. Once again, it is

---

[6]There is some ambiguity in interpreting any conclusion about how heterogeneity influences the level of optimal marginal income tax rates because there are different possible benchmarks for comparison. Ceteris paribus, a higher welfare weight on marginal dollars does favor lower marginal tax rates, as discussed in the text. But, as note 4 mentions, heterogeneity has other effects, including on labor effort. Taking, for example, a case in which the marginal utility of consumption is higher, the analysis in the text indicates that labor effort will also be higher. Hence, for a given observed income level wl, the higher l implies a lower w. Moreover, w is in the denominator of the first-order condition (2.3) for optimal marginal income tax rates, so the lower w favors higher marginal income tax rates, the opposite implication of the higher marginal utility. Accordingly, if labor supply is highly elastic, the effects on optimal tax rates attributable to marginal utility differences could be offset or even reversed.

[7]Realistically, preferences are partially observable, in that age, physical disabilities, family status, or other factors signal preference differences. To an extent, one can interpret the present analysis as applicable to residual heterogeneity within an identifiable class of individuals, although there may be limits on the ability to employ different commodity taxes for different classes. Also note that, in addition to limitations due to feasibility and administrative cost, one can also imagine political constraints, for example, on the use of certain factors such as race that might correlate with preferences.

[8]In addition, in the denominator one would have an integral over types of the product εwf. Regarding w, note that different θ's may induce individuals of a given w to supply different l's, so all those at a given income level wl need not be of the same earning ability w.

helpful to refer to ϕ_k, a generic preference parameter for good k. Differentiating the commodity demands (3.2) with respect to ϕ_k yields

$$(4.1)\quad \frac{dx_k}{d\phi_k} = \frac{\partial x_k}{\partial \phi_k} + \frac{\partial x_k}{\partial \mu}\frac{d\mu}{d\phi_k}.$$

Using expression (3.5) for $d\mu/d\phi_k$, expression (4.1) can be restated as

$$(4.2)\quad \frac{dx_k}{d\phi_k} = \frac{\partial x_k}{\partial \phi_k}\left(1+\frac{q_k\frac{\partial x_k}{\partial \mu}}{w(1-T')\frac{dl}{d\mu}-\sum_{i=1}^{n}\left(q_i\frac{\partial x_i}{\partial \mu}\right)}\right).$$

The denominator of fractional term in large parentheses on the right side of expression (4.2) is the same as the denominator on the right side of expression (3.5) for $d\mu/d\phi_k$, and this was previously shown to be positive. Specifically, the first term of the denominator is positive and each term in the summation is negative. From the latter, it follows that the numerator in the fractional term is negative. Moreover, this numerator is equal to the k^{th} term in the summation in the denominator. Taken together, these features imply that the value of the term in large parentheses is in the interval (0, 1). Therefore, expression (4.2) indicates that the total derivative $dx_k/d\phi_k$ has the same sign as the partial derivative $\partial x_k/\partial\phi_k$ but is smaller in magnitude. This result is in accord with intuition: The total effect on demand is given by starting with the effect when μ, the marginal utility of consumption, is taken to be constant and then dampening it by the resulting adjustment in μ. (For example, if the partial derivative indicates that demand would rise, this rise in demand is financed partly by reducing demands for other goods and partly by increasing labor supply, both of which imply a higher μ, and the higher μ reduces the magnitude of the increase in demand, but not below zero.)

Because each of the partial derivatives $\partial x_k/\partial\phi_k$ was signed in subsection 3.2, in expressions (3.8) through (3.11), we can now consider how preference heterogeneity bears on optimal differentiation of commodity taxes. Suppose initially that the only heterogeneity involves parameter α_i for some good x_i. From expressions (3.8) and (4.2), it follows that individuals who have a higher α_i will have a higher demand for x_i, ceteris paribus. Furthermore, the analysis in subsection 3.2 indicates that the direction of the ideal redistributive adjustment (i.e., the adjustment to T if θ were observable) depends on the SWF since a higher α_i implies a higher marginal utility v_c but also a higher utility level and thus a lower W' if the SWF is strictly concave.[9] For concreteness, suppose that the SWF is utilitarian, in which case W' is constant, so it is optimal to redistribute toward high-α_i individuals. Since they have higher demands for x_i,

[9]Here and throughout this section, one should keep in mind the qualification in note 6 regarding the direction and magnitude of optimal adjustments to marginal income tax rates when those adjustments derive from differences in marginal utilities of consumption.

some degree of subsidy would be optimal.[10] Note that the use of the term subsidy here refers to a relative subsidy (and likewise for later uses of the term tax); instead of subsidizing good x_i, one could instead (equivalently) tax all other goods and adjust T accordingly.

As one generalizes to cases involving preference heterogeneity relating to many commodities, the problem becomes more complex. For example, suppose that there was heterogeneity involving all of the α_i's. The preceding analysis suggests that one would like to relatively subsidize all goods, which is impossible – i.e., the effects of the different subsidies would be offsetting. Taking a more concrete example, suppose that some types have uniformly higher α_i's than do others. Under the utilitarian SWF, it would be ideal to redistribute income to them; however, this cannot be done through commodity taxation because, in this case, relative demands would be unaffected.

Combining the preceding points, it appears that it would be optimal to subsidize goods relatively preferred by types whose overall or average levels of the α_i's are higher. To illustrate, assume that individuals with generally higher α_i's have an α_j that is relatively low for them. The prescription would be to relatively subsidize all goods except good x_j – i.e., to relatively tax good x_j, even if these individuals' α_j's are above the population average (although lower than the average of their own, other α_i's).

Alternatively, suppose that each of the α_i's is independently distributed. In that case, it would seem advantageous under a utilitarian SWF to relatively the subsidize the goods with the greatest variance in the distribution of the α_i's, whereas a tax may be optimal if the SWF is sufficiently concave. This point is most easily seen in the limiting case in which the variance for one of the α_i's is large and that for all of the others approaches zero, which presents the original case in which there is heterogeneity with respect to only one of the α_i's.[11]

Similar analysis applies to the other parameters. For example, higher γ_i's imply both lower marginal utility (see expression 3.10) and higher utility, so we would ideally like to redistribute away from individuals with atypically high γ_i's. Furthermore, high-γ_i individuals will have lower demands for corresponding goods x_i, so it would be ideal to subsidize goods for which some individuals have unusually high γ_i's. (With the α_i's, commodity subsidies favored individuals with high values of the parameter; with the γ_i's, subsidies favor individuals with low values.) But, as with the α_i's, one cannot relatively subsidize all goods, so one must consider relative differences in parameter values.

For the parameter δ, higher values indicate no difference in marginal utility (see expression 3.11) but higher utility and thus a lower W' if the SWF is strictly concave. However, because the value of δ has no influence on commodity demands, commodity taxes cannot be

---

[10]Observe that any effects of such a subsidy across income levels (for example, with a normal good, higher-ability individuals who will earn higher incomes will spend more on the good and thus benefit more from the subsidy, even if they have the same α_i as others) can be offset through the T schedule.

[11]The magnitude of the corresponding ρ_i's would also matter, for ρ_i affects the magnitude of the utility level and marginal utility effects and also the magnitude of demand effects, including the demand elasticity, which is related to the distortionary cost being traded off against the redistributive benefit.

employed directly to accomplish any desired redistribution.

For the parameter β, the situation is more complicated. From expression (3.9), the sign of the marginal utility effect – and, accordingly, from expression (4.2), the sign of the demand effect – depends on the magnitude of the pertinent ρ_i (abstracting from the further adjustment required when the corresponding γ_i does not equal zero). For example, if heterogeneity concerns a particular β_i, and we consider the case in which $\rho_i < 1$, then high-β_i individuals have higher marginal utilities, ceteris paribus, and under a utilitarian SWF it would be optimal to redistribute toward them. In this case, such individuals also have higher demands, so a subsidy would be optimal. If instead $\rho_i > 1$, then high-β_i individuals have lower marginal utilities, so it would be optimal to redistribute away from them. In this alternative case, however, such individuals have lower demands, so again a subsidy (which would favor low-β_i, high-demand individuals) would be optimal. In both instances, note that the ability to use commodity taxation to redistribute is limited the closer ρ_i is to 1; when $\rho_i = 1$, there is no demand effect, so across-preference-type redistribution through commodity taxes is infeasible (and also would not be optimal with a utilitarian SWF, although it would be if the SWF was strictly concave).[12]

The optimal use of differential commodity taxation actually depends on the combination of all of the parameters. As a group, they determine whether it would be optimal to redistribute toward or away from an individual of a given overall preference type θ. Furthermore, as a group they determine the direction of any differences in demand. Even for a specified SWF, such as a utilitarian one, the ideal direction of redistribution sometimes is in the same direction and sometimes is in the opposite direction of the corresponding demand effect. Accordingly, broad generalizations about preference heterogeneity and the signs of optimal deviations from uniform commodity taxation cannot be offered.

It may be possible, however, to identify some particular effects. It seems plausible that some individuals toward whom it would be optimal to redistribute do systematically demand more of certain commodities. This might include those with physical limitations (the goods might be types of medical care or disability accommodations such as wheelchairs) or mental infirmities (the goods might be certain drugs or psychiatric care). To some degree, the analysis of section 3 may be applicable because some of these differences in preferences correspond to observable differences, in which case redistribution should (in the benchmark case with separable labor) be accomplished entirely through the income tax. But some physical and mental infirmities may be more difficult to observe, so differential commodity taxation may play a useful role. In this regard, it should be noted that optimal differential taxation may be appropriate for more than the obvious goods such as those already noted. For example, individuals with more hidden physical disabilities might engage less in physically strenuous activities (skiing) and more in gentler activities; those with certain psychological difficulties likewise may have atypically high demand for some commodities and low demand for others. What matters, it should be recalled, is various individuals' relative demands. Thus, some impaired individuals might have an unusually low α_i for some commodity but that parameter

[12]The existence of little or no demand effect due to differences in β_i's does not, however, imply that there is no distortion in commodity demands from differential taxation.

value may still be relatively high for them, and this parameter value may in turn imply higher demand, which would favor a subsidy.

4.2. Public Goods

To introduce public goods, one can modify the initial utility function to be $u(x, g, l, \theta)$, where g is a vector of $m-n$ public goods. Likewise, the specific utility function (2.4) can be amended as follows:

$$(4.3)\ u(x,l,\alpha,\beta,\gamma,\delta,\rho) = \sum_{i=1}^{n}\left(\alpha_i \frac{(\beta_i x_i + \gamma_i)^{1-\rho_i}}{1-\rho_i} + \delta_i\right) + \sum_{i=n+1}^{m}\left(\alpha_i \frac{(\beta_i g_i + \gamma_i)^{1-\rho_i}}{1-\rho_i} + \delta_i\right) - z(l).$$

(As before, for the case in which $\rho_i = 1$, the fractional expressions are replaced by $\ln(\beta_i x_i + \gamma_i)$ and $\ln(\beta_i g_i + \gamma_i)$ respectively.)

Consider first the case in which θ is observable. With weak labor separability, which expression (4.3) exhibits, the proof in Kaplow (1996) on the optimality of the Samuelson rule for public goods provision – without adjustments for distribution or labor supply distortion – goes through. (The reasoning is analogous to that in subsection 3.1, as the proof in Kaplow (1996) in relevant respects is analogous to that in Kaplow (2006a) for uniform commodity taxation.) Thus, in summing (integrating) individuals' marginal benefits (measured in dollars) from more of a public good, each individual's possibly idiosyncratic valuation would be employed; this sum, whatever it may be, would be compared with the marginal cost of increased provision, just as in the case with homogeneous preferences.[13]

Accordingly, the effect of preference heterogeneity regarding public goods on optimal policy is entirely through the income tax schedules $T(wl, \theta)$, which are customized for each preference type. These effects, however, are qualitatively different from those deriving from heterogeneous preferences for commodities. The reason is that, although differences in the ϕ_k's, for $k > n$ (i.e., for the public goods) affect utility *levels* analogously (utility is increasing in each of α, β, γ, and δ), differences in these ϕ_k's do not affect individuals' *marginal* utilities of consumption. Accordingly, for a utilitarian SWF, there would be no adjustment to the optimal income tax schedule, and for a strictly concave SWF, higher values of these parameters imply higher utility and thus a lower W', so optimal marginal income tax rates would be higher.

This irrelevance result concerning marginal utilities is due to the separability between public and private goods embodied in the utility function (4.3). More generally, there could be

[13]This sum, to be sure, may be influenced by heterogeneity itself. As it turns out, greater heterogeneity (with a constant mean) can raise or lower the sum of marginal benefits, depending on the preference parameter that varies (and the level of the corresponding ρ_i).

marginal utility effects in either direction. For example, better roads may make automobiles more valuable, and one could imagine this arising in a manner that increased individuals' marginal utilities of consumption more for individuals who had a stronger preference for driving. Suppose further that this greater preference operated through the corresponding α_i (perhaps $\alpha_j g_j$ for roads multiplies this α_i). Then high-α_j individuals would have a higher "effective" α_i, which ceteris paribus implies a higher marginal utility of consumption, so under a utilitarian SWF their optimal marginal income tax rates would be lower on this account.

Now consider the case in which θ is unobservable. The analysis of the optimal income tax would be analogous to that described in subsection 4.1 on unobservable heterogeneity in preferences for commodities. It might appear that there is no implication corresponding to the nonuniform commodity tax results since all individuals necessarily receive the same amount of each public good – i.e., differences in consumption patterns do not exist and thus may seem to render differential treatment infeasible. But this supposition is incorrect. By providing more or less of the public good than the amount indicated by the Samuelson rule, different preference types can be favored and disfavored: Raising (lowering) the level of the public good disproportionately benefits those with a higher (lower) marginal utility for that good (assuming distribution-neutral finance of the sort described in subsection 3.1).

The remaining question is which goods should thus be over- or under-provided relative to the level that satisfies the Samuelson rule. Once again, with preferences separable between public and private goods and a utilitarian SWF, there is no basis for deviation since heterogeneity in preferences for public goods does not influence individuals' marginal utilities of consumption. If the SWF is strictly concave or if there are cross-effects, then the previous comments for the case of observable preferences would become relevant.[14] For example, with a strictly concave W, if individuals with overall detrimental preference parameters (i.e., those whose preference parameters generate less utility from public and private goods at a given income level) tend to derive atypically high marginal utility from public good j, then more of that public good should be provided. To be concrete, suppose that individuals who are unusually vulnerable to being mugged also tend to be worse off in other respects (that cannot readily be observed). In this case, greater police protection would have a favorable distributive effect. By contrast, if national parks are most enjoyed by more robust individuals, who tend otherwise to be better off on average, then lower provision would be distributively beneficial.

4.3. Externalities

To examine externalities instead of public goods (considering the two together would be straightforward), one can modify the initial utility function to be $u(x, e, l, \theta)$, where e is a vector of $m-n$ externalities. For concreteness, one might take $m = 2n$ and suppose that each externality is measured by the total consumption of the corresponding commodity. Analogous to expression (4.3), the specific utility function (2.4) becomes:

[14]When there are public goods, optimal commodity taxation may also be influenced in related ways.

$$(4.4)\ u(x,l,\alpha,\beta,\gamma,\delta,\rho) = \sum_{i=1}^{n}\left(\alpha_i \frac{(\beta_i x_i + \gamma_i)^{1-\rho_i}}{1-\rho_i} + \delta_i\right) + \sum_{i=n+1}^{m}\left(\alpha_i \frac{(\beta_i e_i + \gamma_i)^{1-\rho_i}}{1-\rho_i} + \delta_i\right) - z(l).$$

(The modification for the case in which $\rho_i = 1$ is analogous as well.) For a negative externality e_i, the corresponding α_i would be negative. Observe that, in this formulation, externalities have a public good character: Each individual is exposed to the same levels of the e_i's, although different preference types will be affected differently by any particular e_i. Thus, the sum (integral) of marginal external benefits or harms for purposes of applying the Pigouvian rule of setting commodity taxes and subsidies equal to marginal harm and benefit, respectively, is directly analogous to the sum of marginal benefits in the public goods case.

At this point it should be unsurprising that the analysis closely parallels that for public goods. When θ is observable and there is weak labor separability, as in expression (4.4), the proof in Kaplow (2006b) indicating that there should be no adjustment on account of distribution or labor supply distortion to the Pigouvian rule goes through. (So do extensions for partial reforms that are analogous to those considered for commodity taxation.) Once again, heterogeneous preferences would only be relevant in setting the income tax schedules $T(wl, \theta)$, and the pertinent adjustments would follow the same sort of reasoning applicable for public goods. For example, in expression (4.4) in which private goods and externalities have separable effects, externalities – and thus differences in preference parameters regarding externalities – affect utility levels but not the marginal utility of consumption, so heterogeneous preferences regarding externalities would only be relevant to optimal income tax rates to the extent that the SWF is strictly concave. Relaxing this separability assumption would allow for interactions.

When θ is not observable, it may be optimal to depart from the pure Pigouvian rule. Although individuals all experience the same levels of the externalities, by regulating externalities more or less than indicated by the Pigouvian rule one can favor or disfavor different types of individuals.[15] For example, if some pollutant imposes greater marginal harm on more infirm individuals who also tend to be worse off in other (unobservable) respects, greater control of that pollutant would tend to be optimal under a strictly concave SWF. Likewise, just as it may have been optimal to reduce expenditures on national parks that disproportionately benefit more robust individuals who are otherwise better off on average, so it may be optimal to reduce (relative to the Pigouvian optimum) the control of pollution that primarily reduces the enjoyment of wilderness areas.

---

[15]Since the regulation of externalities is taken here to be implemented through commodity tax adjustments, differences in preferences for commodities are also relevant to across-preference-type redistribution, but that is already captured by the analysis in subsection 4.1.

5. Conclusion

Heterogeneous preferences undoubtedly exist and in some instances may be important. Accordingly, it is useful to revisit a variety of optimal policy rules that have been derived in models in which preferences are taken to be homogeneous. The results depend greatly on whether preference differences are observable and on the nature of those differences.

With observability, a number of first-best policy prescriptions in benchmark cases (notably, with weak labor separability) continue to hold. This conclusion is true of the Atkinson-Stiglitz (1976) rule favoring uniform commodity taxation as well as generalizations that do not require the income tax to be optimal and that encompass partial reforms, such as proportional moves toward uniformity. Likewise, results regarding the Pareto optimality of moving public goods provision in the direction indicated by the Samuelson rule and Pigouvian taxes and subsidies toward full internalization of externalities – without regard to concerns for distribution and labor supply distortion – also extend to the present setting.

The characterization of optimal nonlinear income taxation, however, changes in important ways. Preference parameters indicating higher utility levels favor, ceteris paribus, higher marginal income tax rates to the extent that the SWF is strictly concave. But different sorts of preferences parameters have different effects on individuals' marginal utilities of consumption and thus parameters producing higher utility levels may favor higher or lower marginal income tax rates.

When differences in preferences are unobservable, a single income tax schedule must be applied to everyone, and optimal marginal income tax rates are determined similarly to the manner applicable with homogenous preferences. (Roughly, weighted averages substitute for specific values, but the basic formula is the same.) In this case, it is the more specific policy rules that may differ qualitatively from what arises with homogeneous preferences. Differential commodity taxation, public goods provision that deviates from the levels implied by the Samuelson rule, and corrective taxes that over- or under-internalize externalities will be optimal to some degree in cases in which such deviations can indirectly accomplish some of the across-preference-type redistribution that would have been implemented through differentiated income tax schedules if preference differences were observable. Examples were offered involving some possible manifestations of physical and mental disabilities.

Overall, departures from results in models with homogeneous preferences depended on a variety of factors. Different ways in which preferences might vary have qualitatively different effects, even in opposite directions. Furthermore, preference differences interact with each other and with the various policy instruments in complex and subtle ways. The problem of heterogeneous preferences is indeed heterogeneous.

Accordingly, results in existing literature on heterogeneity need to be interpreted as special cases in ways that have not previously been recognized; indeed, somewhat different specifications of preference differences can reverse results. Nevertheless, some plausible conjectures can be offered when it is possible to ascertain the character of preference differences, even if they cannot be observed for each individual. These findings may have direct policy relevance, and they indicate the value of empirical research on preference heterogeneity.

References

Atkinson, Anthony B., and Joseph E. Stiglitz. 1976. "The Design of Tax Structure: Direct Versus Indirect Taxation." *Journal of Public Economics* 6: 55-75.

Atkinson, Anthony B., and Joseph E. Stiglitz. 1980. *Lectures on Public Economics*. New York: McGraw-Hill.

Blackorby, Charles, and David Donaldson. 1988. "Cash Versus Kind, Self-Selection, and Efficient Transfers. *American Economic Review* 78: 691-700.

Boadway, Robin, and Michael Keen. 1993. "Public Goods, Self-Selection and Optimal Income Taxation." *International Economic Review* 34: 463-478

Boadway, Robin, Maurice Marchand, Pierre Pestieau, and Maria Racionero. 2002. "Optimal Redistribution with Heterogeneous Preferences for Leisure. *Journal of Public Economic Theory* 4: 475-498.

Ebert, Udo. 1988. "Optimal Income Taxation: On the Case of Two-Dimensional Populations." University of Bonn, Graduate School of Economics Discussion Paper No. A-169.

Fleurbaey, Marc. 2006. "Is Commodity Taxation Unfair." *Journal of Public Economics* 90: 1765-1787.

Hellwig, Martin F. 2004. "Optimal Income Taxation, Public-Goods Provision and Public-Sector Pricing: A Contribution to the Foundations of Public Economics. Max Planck Institute for Research on Collective Goods Working Paper 2004/14.

Hylland, Aanund, and Richard Zeckhauser. 1979. "Distributional Objectives Should Affect Taxes But Not Program Choice or Design." *Scandinavian Journal of Economics* 81: 264-284.

Kaplow, Louis. 1996. "The Optimal Supply of Public Goods and the Distortionary Cost of Taxation." *National Tax Journal* 49: 513-533.

Kaplow, Louis. 2006a. "On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal." *Journal of Public Economics* 90: 1235-1250.

Kaplow, Louis. 2006b. "Optimal Control of Externalities in the Presence of Income Taxation." NBER Working Paper 12339.

Konishi, Hideo. 1995. "A Pareto-Improving Commodity Tax Reform under a Smooth Nonlinear Income Tax." *Journal of Public Economics* 56: 413-446.

Laroque, Guy. 2005. Indirect Taxation Is Superfluous Under Separability and Taste Homogeneity: A Simple Proof. *Economics Letters* 87: 141-144.

Marchand, Maurice, Pierre Pestieau, and Maria Racionero. 2003. "Optimal Redistribution when Different Workers Are Indistinguishable." *Canadian Journal of Economics* 36: 911-922.

Mirrlees, James A. 1976. "Optimal Tax Theory: A Synthesis." *Journal of Public Economics* 6: 327-358.

Ng, Yew-Kwang. 1984. "Quasi-Pareto Social Improvements." *American Economic Review* 74: 1033-1050.

Saez, Emmanuel. 2002. "The Desirability of Commodity Taxation under Non-linear Income Taxation and Heterogeneous Tastes." *Journal of Public Economics* 83: 217-230.

Sandmo, Agnar. 1993. "Optimal Redistribution When Tastes Differ." *Finanz Archiv* 50: 149-163.

Tarkiainen, Ritva, and Matti Tuomala. 1999. "Optimal Nonlinear Income Taxation with a Two-Dimensional Population: A Computational Approach." *Computational Economics* 13: 1-16.