The Theory of Optimal Taxation

Sørensen, Peter Birch

Publication date:
2009

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Sørensen, P. B. (2009). *The Theory of Optimal Taxation: New Developments and Policy Relevance*. Economic Policy Research Unit. Department of Economics, University of Copenhagen.

EPRU Working Paper Series 2009-09

Economic Policy Research Unit
Department of Economics
University of Copenhagen
Øster Farimagsgade 5, Building 26
DK-1353 Copenhagen K
DENMARK
Tel: (+45) 3532 4411
Fax: (+45) 3532 4444
Web: http://www.econ.ku.dk/epru/

The Theory of Optimal Taxation: New Developments and Policy Relevance

Peter Birch Sørensen

ISSN 0908-7745

April 2009

THE THEORY OF OPTIMAL TAXATION:
NEW DEVELOPMENTS AND POLICY RELEVANCE

by

Peter Birch Sørensen

Abstract: *The theory of optimal taxation has often been criticized for being of little practical policy relevance, due to a lack of robust theoretical results. This paper argues that recent advances in optimal tax theory has made that theory easier to apply and may help to explain some current trends in international tax policy. Covering the taxation of labour income and capital income as well as indirect taxation, the paper also illustrates how some of the key results in optimal tax theory may be derived in a simple, heuristic manner.*

Keywords: Optimal taxation, uniform taxation, tax neutrality
JEL classification: H21, H24, H25

Address for correspondence:
Peter Birch Sørensen
Department of Economics, University of Copenhagen
Studiestraede 6, 1455 Copenhagen K, Denmark
E-mail: pbs@econ.ku.dk

THE THEORY OF OPTIMAL TAXATION:
NEW DEVELOPMENTS AND POLICY RELEVANCE

Peter Birch Sørensen[1]
University of Copenhagen, EPRU and CESifo

In the tradition established by the classical political economists, normative analysis of tax policy tended to follow a principles-oriented approach according to which a good tax system should satisfy certain desirable criteria. For example, Lord Overstone, who served as President of Britain's Royal Statistical Society from 1851 to 1853, thought that a tax should be "productive, computable, divisible, frugal, non-interferent, unannoyant, equal, popular, and uncorruptive" (see the discussion by O'Brien, 2009).

The classical economists rarely discussed the trade-offs between the various goals of tax policy. In particular, they did not pay much attention to the trade-off between redistribution and economic efficiency, since they typically ruled out redistributive progressive taxation as a matter of principle, seeing it as a fundamental threat to property rights. The denouncement of any deviation from proportional taxation was vividly expressed by McCulloch (1845) who argued that "The moment you abandon the cardinal principle of exacting from all individuals the same proportion of their income or of their property, you are at sea without rudder or compass, and there is no amount of injustice and folly you may not commit" (see Creedy, 2009, p. 2).

Following the neoclassical "marginalist" revolution in economic theory, Edgeworth (1897) argued that taxation should involve an equal marginal sacrifice (of utility) for each individual taxpayer in order to minimise the aggregate utility loss imposed on taxpayers. When combined with the neoclassical assumption of declining marginal utility of income, this utilitarian principle of equal marginal (i.e. minimum total) sacrifice did provide a rationale for progressive income taxation.

---

[1] This paper is a revised and extended version of an invited plenary lecture given at the 20th Scientific Meeting of the Società italiana di economia pubblica (SIEP) in Pavia on 25-26 September, 2008. Without implicating her in any remaining shortcomings, I thank my discussant Lisa Grazzini for her comments on my lecture. I am also grateful to Claus Thustrup Kreiner and Etienne Lehmann for valuable comments on an earlier draft of the paper.

Edgeworth was aware that redistributive taxation involves a trade-off between equity and efficiency, but the development of a rigorous coherent framework for analysing this trade-off had to await the seminal work by Mirrlees (1971). However, much of the optimal tax literature building on Mirrleess' contribution has been highly technical and abstract, and for many years this body of theory seemed to offer few robust results. For these reasons many policy makers have tended to see the theory of optimal taxation as being of little practical relevance. In this paper I shall argue that recent advances in optimal tax theory has made that theory easier to apply and may help to explain some current trends in international tax policy. I shall also illustrate how some of the key results in optimal tax theory may be derived in a simple, heuristic manner.

The first part of the paper deals with the theory of optimal taxation of labour income. In Part 2 I focus on optimal indirect taxation, while Part 3 discusses the optimal taxation of income from capital. Part 4 summarises the main conclusions.

1. Optimal taxation of labour income

The Mirrlees model

In the canonical model of optimal income taxation set up by Mirrlees (1971) consumers are assumed to maximise a utility function of the general form

$$U = U\left(C, L\right), \tag{1}$$

subject to the budget constraint

$$C = wL - T\left(wL\right), \tag{2}$$

where C is consumption, L is labour supply, w is the real wage, and $T(wL)$ is a non-linear tax-transfer schedule. The solution to the consumer's problem yields his indirect utility function $V(w)$. In the Mirrlees-model the pre-tax real wage rates are treated as exogenous and taken to reflect the different non-observable ability levels of individual taxpayers. With wage rates being distributed over the interval $\left[\underline{w}, \overline{w}\right]$, $0 \leq \underline{w} < \overline{w} \leq +\infty$, Mirrlees assumed that the benevolent policy maker wishes to maximise an individualistic Bergson-Samuelson welfare function of the form

$$W = \int_{\underline{w}}^{\overline{w}} \Psi\left(V\left(w\right)\right) f\left(w\right) dw, \qquad \Psi' > 0, \quad \Psi'' \leq 0, \tag{3}$$

Here $f(w)$ indicates the density of taxpayers earning the wage rate *w*, and the (numerical) magnitude of the second derivative Ψ'' reflects the strength of the policy maker's preference for equity. The maximisation of (3) takes place subject to the constraint that the government must raise an exogenous amount of revenue *R*:

$$\int_{\underline{w}}^{\overline{w}} T(wL(w))\, f(w)\, dw = R. \tag{4}$$

The solution to the above optimal tax problem is technically demanding and does not yield very clear-cut results regarding the shape of the optimal income tax schedule. To get a feel for the likely shape of this schedule, Mirrlees carried out simulations assuming Cobb-Douglas utility functions, a classical utilitarian social welfare function (with $\Psi' = 1$ and $\Psi'' = 0$) and a log-normal wage distribution. On these assumptions he found that the optimal tax schedule was approximately linear, with an exemption level below which positive net transfers are payable.

Had this early result been robust, it would have had great practical policy relevance, since a linear labour income tax is fairly simple to administer. In particular, because a linear income tax features a constant marginal tax rate, it does not require information on individual incomes, since it can be implemented as a proportional payroll tax combined with a flat transfer to all taxpayers. However, subsequent work by Tuomala (1984) and others revealed that the near-optimality of a linear income tax is not a robust result once one allows for plausible respecifications of utility functions and of the shape of the wage distribution. Atkinson and Stiglitz (1980, ch. 13) also found that the optimal tax schedule deviates substantially from linearity when the social planner has more egalitarian preferences than those implied by classical utilitarianism.

With these discouraging findings, it seemed for a while that optimal tax theory could offer little guidance on income tax design. But building on earlier contributions by Revesz (1989), Piketty (1997), Diamond (1998), and Roberts (2000), Saez (2001, 2002a) showed how a formula for the optimal marginal tax rate at every income level can be derived in terms of the relevant elasticities of taxable income and the properties of the wage distribution. Since these parameters can in principle be observed or estimated empirically, the work of Saez has greatly enhanced the practical applicability of optimal income tax theory. Another important contribution by Saez (2002a) was the explicit allowance for tax distortions to the extensive as well as the intensive margin of labour supply. This is highly relevant since many empirical studies have shown that labour supply is often more elastic on the extensive margin (where workers decide whether or not to participate in the labour market) than on the intensive margin (where people make decisions on their hours of work, given that they are already employed).

The Saez (2002a) model

Given its importance, it is worth restating the key contribution of Saez (2002a) in a heuristic, intuitive manner.[2]

Following Piketty (1997), Saez assumes that employed workers allocate themselves across a range of different occupations involving different levels of effort and income. Workers can decide to participate or to stay outside the labour market. In the latter case they receive the public transfer S_0, whereas a worker employed in job category *i* earns the after-tax income c_i. As the net gain $c_i - S_0$ from employment in occupation *i* increases, more workers move from non-employment into this occupation. This is the extensive labour supply response. Once employed, workers can move one step up or down the job ladder by adjusting their effort. In case of an increase in the gap $c_{i+1} - c_i$ between the net incomes obtainable in job categories *i+i* and *i*, respectively, some workers previously employed in occupation *i* will therefore be induced to upgrade themselves to the higher-earning occupation *i*+1 by increasing their "effort". Similarly, an increase in the net earnings differential $c_i - c_{i-1}$ will spur some workers to upgrade themselves from job category *i*-1 to category *i*. These movements between job categories involving different earnings levels represent the intensive labour supply response. With these assumptions, the fraction h_i of the work force employed in occupation *i* can be specified as

$$h_i = h_i\left(c_i - S_0, c_{i+1} - c_i, c_i - c_{i-1}\right), \tag{5}$$

where $c_i \equiv z_i - T_i\left(z_i\right)$ is the difference between the pre-tax earnings z_i in occupation *i* and the net tax $T\left(z_i\right)$ payable on that income. By assuming that labour supply decisions depend only on the *differences* between the net incomes obtainable in different labour market states, equation (5) implicitly abstracts from income effects (since otherwise labour supply would also depend on the *level* of income). In recent years most empirical studies have indeed tended to find rather small income effects on individual labour supply. Thus there is a good case for considering the benchmark situation with zero income effects.

From (5) we can define the *participation elasticity*

[2] The present heuristic exposition follows the "perturbation method" adopted in the main text of Saez's article but is slightly more elaborate. The rigorous formal proofs of Saez's results may be found in the appendix to his paper.

$$\eta_i \equiv \frac{\partial h_i}{\partial\left(c_i - S_0\right)} \cdot \frac{c_i - S_0}{h_i}, \tag{6}$$

and the *intensive labour supply elasticity*

$$\varsigma_i \equiv \frac{\partial h_i}{\partial\left(c_i - c_{i-1}\right)} \cdot \frac{c_i - c_{i-1}}{h_i}. \tag{7}$$

As shown by Saez (2002a, p. 1070), $\varsigma_i \equiv \frac{\varepsilon_i \cdot z_i}{z_i - z_{i-1}}$, where $\varepsilon_i \equiv \frac{\partial z_i}{\partial\left(1-T'(z_i)\right)} \cdot \frac{\left(1-T'(z_i)\right)}{z_i}$ is the elasticity of taxable income with respect to one minus the marginal tax rate, estimated in numerous recent empirical studies.

Optimal taxation in the Saez model

Consider now the welfare effects of increasing the tax liability by the (small) amount dT_i for all of the occupations $i,\ i+1,\ i+2,...$ up to the highest-earning occupation *J*. Measured in money metric units, the impact on the welfare of an individual in one of these occupations equals the direct impact on disposable income, $-dT_i,$ since any change in disposable income stemming from a change in labour supply has no first-order effect welfare, assuming that the taxpayer has optimised his/her labour supply in the initial equilibrium (this is just an application of the Envelope Theorem). Suppose now that the policy maker's evaluation of the marginal social utility of net income for individuals in occupation j is $g\left(c_j\right).$ The social evaluation of the impact of the tax increase on aggregate private welfare is then given by

$$dW\left(i\right) = -\sum_{j=i}^{J} g\left(c_j\right) h_j dT. \tag{8}$$

Against this private welfare loss one must balance the rise in public revenue. Abstracting from behavioural changes, the direct ("mechanical") revenue gain is

$$dM = \sum_{j=i}^{J} h_j dT. \tag{9}$$

However, public revenue also changes because of the tax-induced changes in labour supply. The tax hike reduces the net income differential $c_i - c_{i-1} \equiv z_i - T_i - \left(z_{i-1} - T_{i-1}\right)$ by the amount dT, whereas it has no impact on the income differential $c_j - c_{j-1}$ between any other two "neighbouring" occupations. With dh_i^i denoting the change in h_i resulting from the *intensive* labour supply response, the revenue loss implied by the change in behaviour on the intensive margin (dB^i) is

$$dB^i = dh_i^i \cdot \left(T_i - T_{i-1}\right) = -\varsigma_i \cdot \frac{h_i dT\left(T_i - T_{i-1}\right)}{c_i - c_{i-1}}, \tag{10}$$

where the last equality in (10) follows from (7). The tax hike also reduces the net income gain $c_j - S_0$ by the amount dT for all persons in job category *i* and above. If dh_i^e indicates the change in h_i stemming from labour supply responses on the *extensive* margin, the drop in revenue occurring as a result of lower labour force participation becomes

$$dB^e = \sum_{j=i}^{J} dh_j^e \cdot \left(T_j + S_0\right) = -dT \sum_{j=i}^{J} \frac{\eta_j h_j \left(T_j + S_0\right)}{c_j - S_0}, \tag{12}$$

where we have used the definition (6) to arrive at the last equality. Under the optimal tax policy, the tax burden on any occupation from *i* and upwards is increased up to the point where the social valuation of the resulting private utility loss is just matched by the government's net revenue gain, i.e. until

$$dW\left(i\right) + dM + dB^i + dB^e = 0. \tag{13}$$

Inserting (8) through (12) into (13), we get the optimal tax rule

$$\frac{T_i - T_{i-1}}{c_i - c_{i-1}} = \frac{1}{\varsigma_i h_i} \sum_{j=i}^{J} h_j \left[1 - g_j - \eta_j \left(\frac{T_j + S_0}{c_j - S_0}\right)\right]. \tag{14}$$

The marginal tax rate at the income level z_i may be defined as

$$m_i \equiv \frac{T_i - T_{i-1}}{z_i - z_{i-1}}, \tag{15}$$

from which it follows that

$$1 - m_i \equiv \frac{\left(z_i - T_i\right) - \left(z_{i-1} - T_{i-1}\right)}{z_i - z_{i-1}} = \frac{c_i - c_{i-1}}{z_i - z_{i-1}} \quad \Rightarrow \quad \frac{m_i}{1 - m_i} = \frac{T_i - T_{i-1}}{c_i - c_{i-1}}. \tag{16}$$

Further, the "participation tax rate" measuring the increase in net taxes imposed when a person moves from non-employment to into employment may be defined as

$$t_j \equiv \frac{T_j + S_0}{z_j}, \tag{17}$$

implying that

$$\frac{T_j + S_0}{c_j - S_0} \equiv \frac{T_j + S_0}{z_j - \left(T_j + S_0\right)} = \frac{t_j}{1 - t_j}. \tag{18}$$

Substituting (16) and (18) into (14), we may write the optimal tax rule as

$$\frac{m_i}{1-m_i} = \frac{1}{\varsigma_i h_i} \sum_{j=i}^{J} h_j \left[1 - g_j - \eta_j \left(\frac{t_j}{1-t_j} \right) \right]. \tag{19}$$

Implications of the optimal tax rule

Equation (19) has a number of implications for tax policy: 1) The optimal marginal tax rate at the income level z_i is lower the higher the intensive labour supply elasticity (ς_i) and the larger the number of taxpayers (h_i) at that income level. This is intuitive, since the efficiency loss from a rise in the marginal tax rate will be greater the more taxpayers who are affected by it and the stronger their labour supply responds to a change in the net gain from additional effort. 2) Since a rise in the marginal tax rate at income level z_i reduces the net labour income of all taxpayers above that earnings level, it induces some of them to exit the labour market. The strength of this extensive labour supply response is larger the higher the participation elasticities η_j, $j = i, i+1, ..., J$, and the resulting loss in net public revenue is greater the higher are the initial participation tax rates t_j of the affected groups and the greater the number of people in these groups. The labour supply response at the extensive margin therefore reduces the optimal marginal tax rate at income level z_i to a larger extent the higher are the values of η_j, t_j and h_j above that income level. 3) Because a higher marginal tax rate at income level z_i cuts into the disposable income of all taxpayers above that level, the optimal marginal tax rate is lower the higher is the social valuation of income for taxpayers above the income level considered; i.e. the larger the values of the welfare weights g_j, $j = i, i+1, ..., J$, and the greater the number of people carrying these weights (h_j).

While the factors mentioned in 1) and 2) reflect how concerns about economic efficiency shape the optimal tax schedule, the parameters in 3) obviously reflect equity concerns. However, note from (19) that allowing for labour supply responses at the extensive margin (an efficiency concern) is equivalent to attaching a higher social welfare weight to groups with high participation elasticities and/or groups with high participation tax rates.

One further point is worth emphasizing: to apply formula (19), no assumptions about individual preferences are needed. The policy maker "only" has to specify his/her distributional value judgements in terms of the social welfare weights g_j and to obtain estimates of the parameters ς_j, η_j, t_j and h_j.[3]

To illustrate the applicability of formula (19), we may consider some instructive special cases. As already mentioned, the analysis above abstracts from income effects on labour supply since such effects have typically been estimated to be small. With zero income effects a lump-sum transfer to an individual taxpayer has no impact on on his/her labour supply and hence does not generate any indirect change in public revenue via this channel. A marginal euro of public funds will then be valued exactly as much as an additional euro distributed evenly across all taxpayers, implying that

$$\sum_{i=0}^{J} h_i g_i = 1. \tag{20}$$

Traditionally the optimal tax literature has abstracted from labour supply responses on the extensive margin, implicitly assuming $\eta_j = 0$. It then follows from (14) that the optimal marginal tax rate at the bottom of the income ladder is

$$\frac{T_1 - T_0}{c_1 - c_0} = \frac{1}{\varsigma_i h_i} \sum_{j=1}^{J} h_j \left(1 - g_j\right) = \frac{1}{\varsigma_i h_i} \left(\sum_{j=0}^{J} h_j - h_0 - \sum_{j=0}^{J} h_j g_j + h_0 g_0 \right), \tag{21}$$

where individuals outside the labour market have been categorized as group zero (indicated by subscript 0), while the lowest-paid workers are categorized as group 1. By definition, we have $\sum_{j=0}^{J} h_j = 1$ which may be inserted into (21) along with (20) to give

$$\frac{T_1 - T_0}{c_1 - c_0} = \frac{\left(g_0 - 1\right) h_0}{\varsigma_1 h_1}. \tag{22}$$

If policy makers have a strong preference for redistribution, the value of the social welfare weight g_0 will tend to be far greater than one, since the weighted average value of g_j across all taxpayers is unity (cf. (20)).[4] According to (22) the marginal tax rate at the bottom of the earnings distribution will then be high. This is a well-known implication of the Mirrleesian model of optimal income taxation when the least productive individuals are outside the labour force. In other words, the

[3] One difficulty is that estimation of the elasticities ς_j and η_j has to be based on the individual labour supplies actually observed whereas formula (19) requires knowledge of the elasticities that would prevail under the *optimal* tax system where labour supply might differ from its current level.

[4] If policy makers believe that people outside the labour market are unemployed mostly because they are "lazy", it is conceivable that the welfare weight g_0 could be smaller than one, but this possibility is disregarded here.

much-debated "poverty traps" caused by the phase-out of social transfers as people belonging to the poorest section of the population move from unemployment into employment are in fact part of an optimal policy when the labour supply of the rest of the work force does not respond at the extensive margin.

However, suppose instead that labour supply responds only at the extensive margin, i.e. $\varsigma_i = 0, \ \forall i$. This alternative benchmark case is of some interest since empirical studies tend to find that labour supply is indeed much more elastic at the extensive than at the intensive margin, at least for the low-skilled and for females (see Heckmann, 1993). With $\varsigma_i = 0$ it follows from (19) that optimal tax policy requires

$$\frac{t_j}{1-t_j} = \frac{1-g_j}{\eta_j}. \tag{23}$$

Thus the optimal participation tax rate for income group *j* is lower the higher the participation elasticity of that group (efficiency concern) and the higher the social valuation of income for members of the group (equity concern). Since $S_0 \equiv -T_0$, we have

$$t_1 \equiv \frac{T_1 + S_0}{z_1} \equiv \frac{T_1 - T_0}{z_1} \quad \Rightarrow \quad \frac{t_1}{1-t_1} = \frac{T_1 - T_0}{z_1 - (T_1 - T_0)} = \frac{T_1 - T_0}{c_1 - c_0}. \tag{24}$$

According to (23) and (24) the optimal marginal tax rate at the bottom is therefore given by

$$\frac{T_1 - T_0}{c_1 - c_0} = \frac{1-g_1}{\eta_1}. \tag{25}$$

Given that the weighted average value of the social welfare weights is unity (see (20)), a policy preference for redistribution will almost surely imply $g_1 > 1$ except in the extreme Rawlsian case where policy makers care only about the poorest group (so that $g_0 = 1$ and $g_j = 0$ for all $j \geq 1$). Thus (25) implies that, when labour supply only responds at the extensive margin, the marginal tax rate on the lowest-paid workers should generally be *negative*. This policy could be implemented by granting a sufficiently large Earned Income Tax Credit which is phased out with rising levels of labour income (to reflect the fact that g_j varies negatively with income). The result in (25) is in stark contrast to the more conventional result reported in (22), and it highlights the importance of allowing for labour supply responses at the extensive margin.

Saez (2002a) applies the general formula (19) to simulate the optimal tax schedule, using data on the U.S. wage distribution plus alternative assumptions about labour supply elasticities and the government's tastes for redistribution. The latter are specified as

$$g_j = g\left(c_j\right) = \frac{1}{p \cdot c_j^{\nu}}, \qquad 0 \leq \nu \leq +\infty, \tag{26}$$

where the parameter ν measures the strength of the preference for redistribution, and p is the marginal value of public funds, calibrated to satisfy (20). For plausible values of the intensive labour supply elasticities, Saez finds that it takes fairly high participation elasticities to rationalize negative marginal tax rates at the bottom, especially if the preference for redistribution is strong. However, with realistic participation elasticities, the lowest-paid workers should face rather low marginal tax rates in order not to discourage their participation, and this could still provide a role for some form of Earned Income Tax Credit (EITC). The new focus in optimal tax theory on the importance of the extensive margin of labour supply thus offers a rationale for the recent trend in many OECD countries towards the introduction of various in-work benefits (such as an EITC) that are intended to "make work pay".

An alternative application of optimal tax theory: deriving implicit social welfare weights

One potential obstacle to the applicability of optimal tax theory is that policy makers may not be able or willing to explicitly specify the social valuation weights g_j in the optimal tax formula (19). However, by "turning the formula on its head", the researcher can ask: what are the magnitudes of the social welfare weights for the various income groups that would make the existing tax system optimal, given realistic assumptions on labour supply elasticities and the distribution of pre-tax earnings? Having estimated the implicit social welfare weights embodied in the current tax-transfer system, the researcher may then confront policy makers and ask them: do these social welfare weights provide a reasonable representation of your actual distributional preferences? If the answer is negative, perhaps policy makers can be induced to reconsider whether the existing tax-transfer system represents a rational trade-off between equity and efficiency. For example, suppose the researcher could point out that the social welfare weights implied by the current tax system are not monotonously decreasing with the taxpayer's level of income. Presumably the policy maker would find it hard to defend such a system and would therefore be willing to consider proposals for reforming it.

To illustrate how such a procedure might work, note that equation (14) may be inverted to give (using $S_0 \equiv -T_0 = c_0$):

$$g_J = 1 - \eta_J \frac{T_J - T_0}{c_J - c_0} - \varsigma_J \frac{T_J - T_{J-1}}{c_J - c_{J-1}}, \tag{27}$$

$$g_i = 1 - \eta_i \frac{T_i - T_0}{c_i - c_0} - \varsigma_i \frac{T_i - T_{i-1}}{c_i - c_{i-1}} + \frac{1}{h_i} \sum_{j=i+1}^{J} h_j \left[1 - g_j - \eta_j \left(\frac{T_j - T_0}{c_j - c_0} \right) \right]. \tag{28}$$

Equation (27) may be used to estimate the implicit social welfare weight for the top income group (group *J*), given estimates of the labour supply elasticities η_J and ς_J plus data on the net tax payments T_J, T_{J-1} and T_0 and the disposable incomes c_J, c_{J-1} and c_0. Once this has been done, one can apply (28) to calculate the implicit welfare weights for groups *J*-1, *J*-2,…etc. in a recursive manner, all the way down to group 1, using data on incomes and tax payments plus estimated elasticities. The welfare weight g_0 may then finally be calculated from (20).

Such an exercise has recently been undertaken by Spadaro (2008) for seven western European welfare states. Spadaro uses data from the tax-benefit calculator EUROMOD which groups the population into ten income deciles. Interestingly, he finds that the implicit social welfare weights are non-monotonic in all the countries considered.[5]

Optimal taxation in imperfect labour markets

The bulk of the literature on optimal income taxation abstracts from labour market distortions other than those caused by the tax system. In competitive labour markets, a switch from proportional to progressive income taxation increases the deadweight loss from taxation, except under highly implausible assumptions regarding the income effects of taxation (see Sandmo, 1983). But once one allows for non-tax distortions due to labour market imperfections, some amount of tax progressivity can actually be defended on pure efficiency grounds (see Sørensen, 1999). For example, in unionised labour markets union wage setting tends to generate involuntary unemployment. If the government raises the marginal tax rate while holding the average tax rate constant (e.g. by raising the personal exemption level), unions can be motivated to moderate their wage claims, thus paving

[5] Implicit social welfare weights may also be derived from the condition that the marginal social cost of raising an additional euro of tax revenue should be equalized across all income groups under the optimal tax system. Applying this method to Danish data, Petersen (2007) likewise finds that the Danish tax-transfer system implies non-monotonic social welfare weights, assigning higher welfare weights to the middle income groups than to the poorest groups.

the way for lower unemployment. The reason is that, with a higher marginal tax rate, it becomes less costly for unions to “buy” more jobs through wage moderation, since a given fall in the pre-tax wage rate will now lead to a smaller drop in the after-tax wage. Stronger tax progressivity also generates wage moderation and lower involuntary unemployment in efficiency wage models where employers pay wages above the market-clearing level as a means of inducing higher productivity of their workers. The explanation is that a rise in the marginal tax rate reduces the effectiveness of a high (pre-tax) wage rate as an instrument for encouraging high labour productivity, given that workers care about after-tax rather than pre-tax wages.

These observations do not imply that tax progressivity is a “free lunch” in imperfect labour markets, since higher marginal tax rates also have distortionary effects, e.g. by inducing unions to bargain for fewer working hours, and by reducing work efforts in an efficiency wage setting. However, even when these tax distortions are accounted for, the numerical analysis in Sørensen (1999) suggests that a substantial degree of tax progressivity can be rationalised on pure efficiency grounds, especially when unemployment benefits are generous.

A challenging task is to analyse optimal taxation when the tax system must serve the goal of redistributing income while at the same time accounting for non-tax labour market frictions. This problem was recently addressed by Boone and Bovenberg (2004) and Hungerbühler et al. (2006) for an economy with labour market frictions stemming from imperfect information that gives rise to job search. When an employer with a job vacancy has been matched with an unemployed job seeker, the wage bargain between the two parties determines the distribution of the rent from the job match.[6] The worker’s share of the rent represents the return to his job search effort, and the employer’s share is the return to his investment in searching for a worker. To generate an efficient labour market equilibrium, the employer’s share of the surplus from the match should correspond to the increase in the probability of a match occurring when he posts an extra vacancy. This efficiency condition is known as the Hosios condition (Hosios, 1990). However, if the worker’s bargaining power is “too” strong (weak), his share of the rent will be larger (smaller) than the efficient share, thus weakening (strengthening) the incentive for employers to post vacancies and thereby generating too much (little) labour market slack and hence too much (little) unemployment.

Hungerbühler et al. (2006) analyse the optimal non-linear tax-transfer system in the benchmark case where the Hosios condition is met so that the labour market equilibrium would be

[6] The rent arises from the fact that neither party can immediately and costlessly find another match.

efficient in the absence of tax. The introduction of a positive marginal labour income tax rate for redistributive purposes induces firms and workers to bargain for lower gross wages, because it raises the cost to the employer of providing the worker with any given increase in the after-tax wage while at the same time reducing the cost to the worker of conceding more profit to the employer by accepting a lower pre-tax wage rate. Hence the optimal tax system involves a level of wages and unemployment *below* the efficient level when the Hosios condition is met, since it is optimal for the government to accept this labour market distortion in return for the redistributive gain from progressive taxation. Indeed, since progressivity implies that the average tax rate increases with income, and since a rise in the average (as opposed to the marginal) tax rate generates upward wage pressure in the Nash bargaining set-up considered by Hungerbühler et al., they find that it is optimal to set marginal tax rates above the average tax rates to moderate wages. Simulating their model for plausible parameter values, they show that the optimal marginal tax rates in their economy with job search frictions tend to be substantially higher than the optimal tax rates in the traditional Mirrleesian optimal tax model with a competitive labour market.

The study of optimal taxation in imperfect labour market is still in its infancy, but the work discussed in this section suggests that it may generate new interesting insights with important implications for tax policy.

2. Optimal indirect taxation

Does optimal tax theory offer any useful guidance for indirect taxation? This is the topic addressed in the present part of the paper.

It is well known that a uniform indirect ad valorem tax on all goods and services is equivalent to a proportional tax on labour income.[7] From a theoretical perspective, the main issue is thus whether there is a case for introducing a system of differentiated commodity taxes? And if so, which commodities should bear the highest (lowest) rates of tax? There are two strands of optimal tax literature dealing with this issue.

The Ramsey approach to indirect taxation

The first one, building on the classical contribution by Ramsey (1927), assumes that the government has to raise some given amount of revenue via indirect taxes. When one abstracts from consumer heterogeneity, the optimal tax problem boils down to raising the required amount of revenue in a manner that minimises the total deadweight loss. This leads to the famous Ramsey rule which says that indirect taxes should be designed so as to cause an equi-proportionate reduction in the compensated demands for all commodities. Thus the optimal indirect tax system seeks to avoid distorting the quantitative pattern of consumption, but since the own price and cross price elasticities will generally differ across commodities, commodity taxes (as a percentage of the total consumer price) should generally be differentiated to induce the same relative reduction in all quantities demanded.

---

[7] We can be a little more precise here. Suppose the consumer's life cycle is divided into two periods, indicated by subscripts 1 and 2. Let C_i, W_i, Π_i and T_i denote consumption, wage income, rent income ('pure profit') and transfer income in period *i*, respectively, and suppose the consumer receives an inheritance *I* in period 1 and leaves a bequest *B* at the end of period 2. If *S* is the saving undertaken in period 1, *r* is the interest rate, and *t* is the uniform indirect tax on consumption, assumed constant over time, the consumer's budget constraints in the two periods are:

$$\text{Period 1:} \quad S = W_1 + \Pi_1 + T_1 + I - (1+t)C_1 \,, \qquad \text{Period 2:} \quad (1+t)C_2 + B = (1+r)S + W_2 + \Pi_2 + T_2 \,.$$

Eliminating *S* and consolidating, one obtains the lifetime budget constraint:

$$C_1 + \frac{C_2}{1+r} = (1-\tau)\left(W_1 + \Pi_1 + T_1 + \frac{W_2 + \Pi_2 + T}{1+r} + I - \frac{B}{1+r} \right), \qquad \tau \equiv \frac{t}{1+t} \,.$$

This shows that a uniform ad valorem commodity tax levied at the rate *t* is equivalent to a proportional tax levied at the rate $\tau = t/(1+t)$ on the sum of wages, rents, transfers and the present value of net bequests received $\left(I - \frac{B}{1+r}\right)$.

The well-known inverse elasticity rule – stating that the optimal commodity tax rates are inversely related to the (compensated) own-price elasticity of demand – is a special case of the Ramsey rule, holding only when the cross price elasticities of demand for the taxed commodities are zero. The inverse elasticity rule essentially seeks to minimise tax distortions to labour supply, for when cross price elasticities in commodity demands are zero, the only way a tax-induced rise in the consumer price can reduce the demand for some commodity is by causing substitution from material consumption towards leisure. In such a setting a low own price elasticity of demand means that a commodity tax has little discouraging effect on labour supply.

The optimal tax rule discovered by Corlett and Hague (1953) is another special case of the Ramsey rule derived in a context with only two taxed commodities plus (untaxed) leisure. The Corlett-Hague rule states that the commodity which is more complementary to (less substitutable for) leisure should carry a relatively high tax rate to offset the tendency of the tax system to induce substitution towards leisure. Thus uniform taxation is optimal only in the special case where both commodities are equally substitutable for (complementary to) leisure. Again, we see that optimal indirect tax design seeks to minimise the tax distortions to labour supply that inevitably occur when only commodities (but not leisure) can be taxed.

More generally, if utility is generated by consumption of the goods bundle $x_0, x_1, x_2, ..., x_n$, where x_0 is untaxed leisure, it can be shown (see Sandmo (1974) or Sadka (1977)) that uniform commodity taxation is optimal if the utility function takes the form

$$U\left(x_0, x_1, ..., x_n\right) = U\left(x_0, v\left(x_1, ..., x_n\right)\right) \tag{29}$$

where the subutility function $v\left(x_1, ..., x_n\right)$ is homothetic. In other words, uniform commodity taxation is optimal if preferences are separable in leisure and commodities – so that all commodities are equally substitutable for leisure - and if all commodities have the same income elasticity of demand. The intuition is that when income elasticities are identical, the equi-proportionate reduction of all compensated commodity demands prescribed by the Ramsey rule also requires the same relative reduction of the *uncompensated* demands for all commodities. Since a uniform ad-valorem tax on all commodities is equivalent to a proportional labour income tax, it will indeed generate the same relative fall in the consumption of all commodities when they all have the same income elasticity and are all equally substitutable for leisure (labour).

The classical Ramsey rule focuses on the pure efficiency aspects of indirect taxation by abstracting from consumer heterogeneity. Diamond (1975) showed how the Ramsey rule is modified in a world of heterogeneous consumers where policy makers trade off efficiency against their redistributional goals. His analysis indicates that while efficiency may call for relatively high commodity tax rates on leisure complements, concerns about equity call for relatively low tax rates on commodities that weigh heavily in the budgets of low-income families.

The Mirrleesian approach to indirect taxation and the Atkinson-Stiglitz theorem

Although Diamond (op.cit.) introduced equity concerns into the theory of optimal indirect taxation, he did not account for the role that a non-linear income tax could play as a means of achieving the policy maker's distributional goals. Another strand of literature on indirect taxation – which might be termed "the Mirrleesian approach" – asks the question: which (if any) commodity taxes should supplement the income tax in order to attain an optimal trade-off between equity and efficiency when the government has to raise a given amount of total revenue?

This literature assumes that consumers have different abilities, reflected in their wage rates. The government cannot observe individual wage rates, but it observes individual incomes and is hence able to levy a non-linear income tax. At the same time, the government does not observe the individual taxpayer's consumption of a particular good, so commodity taxes must be impersonal and hence linear. Note that since uniform commodity taxation is equivalent to a proportional wage income tax, the issue for indirect tax policy is whether there is any need for differentiated commodity taxes when the government can levy a personal tax on labour income.

The benchmark result in this line of research was established early on by Atkinson and Stiglitz (1976) who showed that if preferences are weakly separable in leisure and all other goods taken together, it is inoptimal to differentiate taxes across commodities when the government optimises the non-linear labour income tax. Weak separability of preferences implies that the utility function takes the form (29), but the Atkinson-Stiglitz theorem does *not* require that the subutility function $v\left(x_1,......,x_n\right)$ be homothetic.

The intuition behind the Atkinson-Stiglitz theorem may be explained as follows. The nonlinear income tax is deployed to achieve the optimal amount of redistribution. If the tax system becomes too progressive (relative to the optimum), an individual with a high wage rate will choose to work less so as to "mimic" the income level of an individual with a lower wage rate. In that case the two persons will pay the same amount of income tax and have the same disposable income, but the person with the higher wage rate will enjoy more leisure. However, the government obviously cannot use the income tax to achieve a further redistribution of welfare from high-ability to low-ability individuals when the former persons mimic the incomes of the latter. If differentiated commodity taxes could reduce the incentive for mimicking, the scope for redistribution via the income tax would increase. But with separability between goods and leisure, the government cannot use differentiated commodity taxes to impose a higher tax burden on high-ability persons by exploiting any relationship between leisure and the consumption of taxable commodities, since the high-ability individuals choose the same commodity bundle as the low-ability persons whom they mimic, even though they consume more leisure. Hence differentiated commodity taxes cannot play any useful role by relaxing the non-mimicking constraints that limit the government's ability to redistribute income. Differentiated taxation will only introduce distortions in commodity choices and hence it is not optimal.

Atkinson and Stiglitz proved their theorem by assuming that the government optimises the non-linear income tax schedule. More recently, Laroque (2005) and Kaplow (2006) have shown that as long as preferences are weakly separable in leisure and commodities, differentiated commodity taxes are undesirable even when the non-linear income tax schedule is *not* optimal. Since this strong result is not so well known, and since Kaplow's proof strengthens the intuition for the Atkinson-Stiglitz theorem, it is worth providing a sketch of his proof.[8] Suppose each consumer has a utility function of the separable form (29), with leisure x_0 being equal to $E-L$, where *E* is the time endowment and *L* is labour time. An individual with the pre-tax wage rate *w* (reflecting his exogenous productivity) maximises the utility function (29) subject to the budget constraint

$$\sum_{i=1}^{n}\left(p_i+t_i\right)x_i = wL - T\left(wL\right), \tag{30}$$

where p_i is the (fixed) producer price of commodity *i*, t_i is the tax rate on that commodity, and $T\left(wL\right)$ is the consumer's labour income tax bill, given the non-linear tax schedule $T\left(\bullet\right)$. With $\boldsymbol{t}$ denoting the vector of commodity tax rates, we may now introduce the indirect subutility function $V\left(t,T\left(wL\right),wL\right)$, defined as the value of subutility $v\left(x_1,......,x_n\right)$ maximised over all the commodities $x_1,....,x_n$, where the commodity tax vector $\boldsymbol{t}$, the income tax schedule $T\left(wL\right)$, and pre-tax income *wL* are taken as given. Note that the indirect subutility depends only on income, *wL*, but not on labour supply *L*. This is a consequence of the separable utility function (29) which

[8] The sketch offered here is a condensed version of the heuristic exposition in Kaplow (2008, pp. 125-133). The rigorous mathematical proof is given in Kaplow (2006). Working independently of Kaplow, Laroque (2005) also provided a rigorous proof of the redundancy of indirect taxation under the assumptions stated above.

implies that the marginal rate of substitution between any two commodities is independent of the amount of labour.

Starting from an initial situation with differentiated commodity taxation (different ad valorem tax rates) where there exists i, j such that $(p_i + t_i)/(p_j + t_j) \neq p_i / p_j$, consider now a move towards a uniform commodity tax vector ***t**** satisfying $(p_i + t_i^*)/(p_j + t_j^*) = p_i / p_j$ for all i, j. By going through the following steps, Kaplow (2006) shows that this reform allows the government to generate a Pareto improvement. First, suppose the government introduces a new nonlinear income tax schedule $T^*(wL)$ which has the property that, at all income levels, the utility of taxpayers will be the same after the reform as before the reform, provided they do not change their labour supply. If the pre-reform tax schedule is $T(wL)$, the post-reform tax schedule will thus satisfy $V(t, T(wL), wL) = V(t^*, T^*(wL), wL)$ for all *wL*, since in this case all taxpayers will then also enjoy the same total utility $U(E - L, V(\bullet))$ before and after the reform, assuming that their labour supply L does not change. And in fact no taxpayer will have any incentive to change his/her labour supply. The reason is that the adjustment of the income tax schedule is undertaken for all income levels, so when a taxpayer with wage rate *w* varies his labour supply, thereby varying *wL*, he will find that the equality $V(t, T(wL), wL) = V(t^*, T^*(wL), wL)$ holds for all choices of *L*. Therefore, if L^o was the optimal labour supply supply that maximised total utility $U(E - L, V(\bullet))$ before the reform, it must still be the optimal labour supply after the reform, given that the maximum attainable subutility $V(\bullet)$ is unchanged for any choice of *L*.

The next step in the proof is to show that the move towards uniform commodity taxation combined with the utility-preserving change in the income tax schedule will generate an increase in government revenue. Suppose for a moment that taxpayers do not respond to the change in relative commodity prices, continuing to consume the same amounts of each individual commodity as before. With an unchanged labour supply in all income groups, the utility-preserving change in the income tax schedule would then imply that the change in indirect tax payments would be exactly offset by the change in income tax payments so as to leave real disposable incomes unchanged at all income levels. As a consequence, total tax revenue would also be unchanged. In reality, consumers would of course react to the commodity tax reform and the ensuing change in relative consumer prices by substituting towards those commodities that have become relatively cheaper. Since this substitution can only make consumers better off, it allows a further increase in the income tax bill at each income level without reducing consumer utility below the pre-reform level. At the same time, the substitution across commodities does not change the total revenue from indirect taxation, since all commodities bear the same ad valorem tax rate after the reform. The utility-preserving reform of direct and indirect taxation must therefore increase total public revenue, allowing the government to create a Pareto improvement by distributing the extra revenue to all taxpayers.

Although popular and important as a benchmark case, the assumption that preferences are separable in leisure and commodities is of course rather restrictive. Christiansen (1984) analysed which commodity taxes should supplement an optimal nonlinear income tax when preferences are not separable. He found that a commodity should be taxed (subsidised) if it is positively (negatively) related to leisure in the sense that more (less) of the good is consumed if more leisure is obtained at constant income. This result has the same flavour as the Corlett-Hague rule: the indirect tax system should discourage the purchase of commodities that tend to be consumed jointly with leisure. The intuition for Christiansen's result follows directly from the intuition underlying the Atkinson-Stiglitz theorem: by taxing complements to leisure, the government makes it less attractive for the more productive individuals (who must sacrifice less leisure to earn a given amount of income) to mimic the income of the less productive individuals, so in this way commodity taxes relax the incentive-compatibility constraints that restrict the government's ability to redistribute income via the income tax. Saez (2002b) has extended Christiansen's analysis to a setting with heterogeneous consumer tastes, showing that the optimal nonlinear labour income tax should be supplemented not only by excises on commodities that are consumed jointly with leisure, but also by taxes on commodities for which high-income earners tend to have a relatively strong taste.

Further arguments for uniform indirect taxation

The empirical work by Browning and Meghir (1991) as well as casual observation strongly indicates that some commodities are better substitutes for leisure than others, i.e. consumer preferences are in fact not weakly separable in leisure and commodities. This would seem to call for a system of differentiated commodity taxes with the characteristics suggested by Christiansen (1984) and Saez (2002b). A number of practical arguments nevertheless speak in favour of uniform taxation.

First of all, the government may simply lack the solid evidence on the compensated cross-price elasticities with leisure that would be needed to implement the optimal differentiation of commodity taxes. Second, a commodity tax system differentiated according to the principles of optimal tax theory would require frequent changes in tax rates in response to changes in tastes and technologies. This would introduce a potentially welfare-reducing element of risk and uncertainty into the tax system. Third, a uniform VAT is easier to administer and less susceptible to fraud than a VAT system with several differentiated rates, since a uniform VAT does not require any borderlines to be drawn between different categories of goods. Fourth, acceptance of differentiated taxation as a general principle might invite special interest groups to lobby for low tax rates on particular economic activities, so adherence to a principle of uniformity may provide a stronger bulwark against wasteful lobbyism.

All of this suggests that uniformity should be the guiding principle for indirect taxation, except where consumption of specific commodities generates obvious externalities that need to be corrected through high excises.[9] However, recent contributions to optimal tax theory by Kleven, Richter and Sørensen (2000)[10] and Kleven (2004), allowing for home production along with market production, have pointed to some specific areas where deviations from uniform taxation may be warranted.

Optimal indirect taxation with household production: a simple example

To illustrate the importance of home production for optimal taxation, suppose the representative consumer produces services within the household (S^h) subject to the concave household production function

$$S^h = h(H), \qquad h' > 0, \quad h'' < 0, \tag{31}$$

where H denotes hours spent working in the home. Suppose further that services can also be purchased in the market place so that total service consumption (S) is the sum of services bought in the market (S^m) and services produced at home:

$$S = S^m + S^h. \tag{32}$$

If the consumer spends L hours working in the market and his total time endowment is E=1, his consumption of leisure (ℓ) will be

$$\ell = 1 - L - H. \tag{33}$$

[9] Sandmo (1975) provides the classical treatment of optimal commodity taxation in the presence of externalities.
[10] The paper by Kleven et al. (2000) built on the earlier work by Sandmo (1990).

Apart from leisure and services, the consumer also consumes "goods" (*G*). For concreteness, let utility be given by

$$U = U\left(\ell, v\left(G,S\right)\right) = \ell^{\alpha}\left(G^{\beta}S^{1-\beta}\right)^{1-\alpha}, \qquad 0<\alpha<1, \quad 0<\beta<1. \tag{34}$$

Thus preferences are weakly separable in leisure and commodities, and the sub-utility function $G^{\beta}S^{1-\beta}$ is homothetic. As mentioned earlier, with such preferences uniform commodity taxation would be optimal in the absence of home production, but as we shall see, this conclusion no longer holds when home production is allowed for.

Without loss of generality we may assume that it takes one unit of market work to produce one unit of each of the two commodities *G* and *S*. Choosing leisure as the numeraire good (i.e. normalising the wage rate at unity) and assuming that labour is the only factor of production, the producer prices of *G* and *S* will then both be equal to one. With t_G and t_S denoting the unit commodity taxes imposed on *G* and *S*, respectively, the consumer's budget constraint therefore becomes

$$P_G G + P_S S^m = L, \qquad P_G = 1+t_G, \quad P_S = 1+t_S, \tag{35}$$

where P_G and P_S are consumer prices, and $L\ \left(=wL\right)$ is the consumer's market income. Using (31) through (33), we may rewrite (35) as

$$P_G G + P_S S + \ell = Y, \qquad Y \equiv 1 + P_S h\left(H\right) - H. \tag{36}$$

The variable *Y* is the consumer's "full income", consisting of his potential market income (=1) plus the "profit" from home production, $P_S h\left(H\right) - H$. According to (36) this full income may be spent on acquiring commodities or on "buying" leisure (by abstaining from work).

The consumer maximises utility (34) subject to the budget constraint (36). The solution to this problem can be shown to imply that

$$\ell = \alpha Y, \qquad P_G G = \beta\left(1-\alpha\right)Y, \qquad P_S S = \left(1-\beta\right)\left(1-\alpha\right)Y, \tag{37}$$

$$P_S h'\left(H\right) = 1 \quad \Rightarrow \quad H = H\left(t_S\right), \quad H' = -\tfrac{h'}{P_S h''} > 0. \tag{38}$$

According to (38) the consumer engages in home production until the resulting marginal saving on services bought in the market, $P_S h'\left(H\right)$, equals the marginal opportunity cost of working at home rather than in the market $(= w = 1)$. Since $P_S = 1+t_S$, we see that this behaviour implies that the time spent on home production rises with the tax rate imposed on services delivered from the market. Using the results in (38) and recalling that $S^h = h(H)$, we note that full income depends on t_S in the following way:

$$Y(t_S) \equiv 1 + (1+t_S)h(H(t_S)) - H(t_S), \qquad Y' = S^h. \tag{39}$$

Inserting the solutions in (37) into the direct utility function (34), we obtain the indirect utility function

$$V = c \cdot \left[(1+t_G)^{\beta} (1+t_S)^{1-\beta} \right]^{\alpha-1} \cdot Y(t_S), \tag{40}$$

where *c* is a constant depending solely on the taste parameters α and β. We may now prove that, starting from an initial situation with uniform taxation ($t_G = t_S$), the government can increase welfare by moving towards a situation where $t_G > t_S$. We do so by showing that if the government raises t_G and lowers t_S in a way that maintains constant utility, it will earn additional revenue. Clearly the government will then be able to raise consumer welfare by transferring the additional revenue back to the consumer. Differentiating (40) and recalling from (39) that $Y' = S^h$, one finds that

$$dV = 0 \quad \Rightarrow \quad dt_G = \left(\frac{P_S S^h - (1-\beta)(1-\alpha)Y}{\beta(1-\alpha)Y} \right) \cdot dt_S. \tag{41}$$

Note that since $S = S^m + S^h$ and $(1-\beta)(1-\alpha)Y = P_S S$ (cf. (37)), the numerator on the right-hand side of (41) equals $-P_S S^m$, so as long as the consumer buys some services from the market ($S^m > 0$), the maintenance of a constant utility level will indeed require a rise in t_G when t_S is lowered.

Consider now how this utility-preserving perturbation of tax rates will affect government revenue. Using (32) and inserting the consumer's optimal demands for *G* and *S* stated in (37), we may write total revenue as

$$R = t_G G + t_S S^m = t_G G + t_S (S - S^h) = \left[\beta \left(\frac{t_G}{1+t_G} \right) + (1-\beta) \left(\frac{t_S}{1+t_S} \right) \right] Y(t_S) - t_S h(H(t_S)), \tag{42}$$

which may be differentiated to give

$$dR = \frac{\beta(1-\alpha)Y}{(1+t)^2} \cdot dt_G + \left\{ \frac{(1-\beta)(1-\alpha)Y}{(1+t)^2} + \left[(1-\alpha) \left(\frac{t}{1+t} \right) - 1 \right] S^h - \left(\frac{t}{1+t} \right) H' \right\} \cdot dt_S, \tag{43}$$

where we have used (31), (38), (39) and the fact that $t_G = t_S = t$ initially. Substituting (41) into (43), we obtain:

$$dR = -\left(\frac{t}{1+t}\right)\left[\alpha S^h + H'\left(t_S\right)\right]\cdot dt_S > 0 \quad \text{for} \quad dt_S < 0. \tag{44}$$

Thus government revenue does in fact increase, enabling the government to raise consumer welfare by recycling the extra revenue.[11] Note that in the absence of home production we would have $S^h = H' = 0$, so in that case there would be no revenue gain and hence no welfare gain by deviating from uniform taxation, according to (44). Absent home production, our model thus reproduces the standard result that uniform commodity taxation is optimal when preferences are separable in leisure and commodities and utility is homothetic in commodities. But once home production is allowed for, it follows from our analysis that commodities which can be produced at home as well as in the market economy should be taxed more lightly than commodities which cannot be produced within the household.

The practical importance of this result is that it is easy to think of commodities that would be candidates for reduced taxation under this principle. For example, housing repair and repair of other consumer durables, child care, cleaning and window-cleaning, garden care, cooking etc. are all consumer services that can either be produced at home or be delivered from the market. According to the analysis above such services should be taxed more lightly than, say, manufactured goods that cannot realistically be produced within the household.

Kleven, Richter and Sørensen (2000) show that this conclusion will almost surely hold also when preferences are not separable and homothetic. In fact they find that even if the consumer services which can be produced at home as well as in the market might be complementary to leisure, it may still be optimal to impose a relatively low tax rate on these services. The intuition for this modification to the Corlett-Hague rule is that a high tax on complements to leisure is not an efficient way of stimulating tax-discouraged labour supply to the market if such a tax induces a shift from market production to home production. The point is that the optimal tax system must minimise the distortionary substitution away from market activities towards untaxed activities. Taxes should distort the pattern of *market* activity as little as possible, and this calls for lenient taxation of those

[11] There are two reasons for the rise in revenue. First, even if home production were unchanged, the fall in t_s reduces the revenue loss caused by the fact that services produced in the home cannot be taxed. This effect is captured by the first term in the square bracket in (44). In addition, the fall in t_s induces consumers to substitute market-produced services for home production, thereby increasing the tax base. This is captured by the last term in the square bracket in (44).

market activities that can most easily be replaced by home production. These activities would typically include the consumer services mentioned above.

The analysis above uses a conventional specification of utility which abstracts from the fact that consumption is a time-consuming activity. Kleven (2004) analyses optimal commodity taxation in the generalised household production framework proposed by Becker (1965) where all utility-generating consumption activities require the combination of some good or service with household time spent on the act of consumption (or on acquiring the commodity).[12] In this framework households combine a "commodity input" (the purchase of some good or service) with an input of time into utility-generating household activities (termed "household production"). Kleven (op.cit.) shows that the optimal commodity tax system imposes relatively high tax rates on commodities whose consumption require a large input of household time. In this way the optimal tax system minimises the amount of time that is diverted from market work to consumption activity within the household sector. Indeed, when all utility-generating activities require a positive input of commodities as well as time, and when all commodities can be taxed, Kleven finds that the optimal commodity tax rate on some commodity *j* depends only (and inversely) on the ratio of the value of the commodity input to the sum of the values of the commodity input and the time input in the consumption activity in which commodity *j* is used. In other words, it is not necessary for policy makers to know all the compensated own price and cross price elasticities of demand to implement the optimal tax system. In principle, all that is needed is a combined survey of consumption expenditures and household time allocation. This information requirement is clearly less daunting than the information on compensated cross price elasticities with leisure needed to implement the classical Ramsey tax rule.

According to Kleven's analysis any type of consumption activity which requires little time, or even saves time, should carry a relatively low tax rate. The type of consumer services mentioned earlier in this section typically have this property: hiring somebody in the market to supply a service rather than engaging in do-it-yourself activities saves household time, so such services should be favoured by the tax system, just as implied by the analysis based on the more traditional model of household production presented above.

It is interesting to note that several European countries have experimented with reduced rates of tax or direct subsidies to a number of labour-intensive services that are easily substitutable for home-produced services, and the EU has recently allowed its member states to

[12] Boadway and Gahvari (2006) also analyse optimal commodity taxation when consumption is time-consuming.

apply reduced rates of VAT on such services. The contributions to optimal tax theory discussed in this section suggest that there is indeed a rationale for this policy, and the numerical simulations in Sørensen (1997) indicate that the efficiency gains from reduced tax rates on this type of services could be significant.

3. Optimal taxation of income from capital

Should capital income be taxed?

The analysis of indirect taxation has important implications for the long-standing debate on the taxation of income from capital. The fundamental issues in this debate are whether capital income should be taxed at all, and if so, whether all returns to capital should be taxed at the same rate, i.e. whether capital income taxation should be "neutral"?

The Atkinson-Stiglitz theorem discussed in Part 2 provides a useful starting point for addressing the first issue: should we tax income from capital? Consider the standard life cycle model where the consumer supplies labour *L* during young age and enjoys total consumption C_1 and C_2 during young and old age, respectively. Suppose that labour income is subject to the nonlinear tax schedule $T(wL)$, where *w* is the pre-tax wage rate, and assume further that the return to capital (*r*) is taxed at the rate t_r. Denoting savings during young age by *S*, we may then write the consumer's budget constraints as

$$S = wL - T(wL) - C_1, \qquad C_2 = \left[1 + r(1-t_r)\right]S,$$
$$\Downarrow \tag{45}$$
$$C_1 + \frac{C_2}{1+r(1-t_r)} = wL - T(wL) \;\Leftrightarrow\; C_1 + \left[\frac{1}{1+r} + \frac{rt_r}{(1+r)(1+r(1-t_r))}\right]C_2 = wL - T(wL).$$

As shown by the second term in the square bracket in (45), a capital income tax may be seen as an excise tax on future consumption, since it raises the relative price of future consumption above the level $1/(1+r)$ that would prevail in the absence of capital income taxation. If we normalize the consumer's time endowment at unity, his consumption of leisure during young age is 1-*L*. Suppose now that his lifetime utility is given by

$$U = U(1-L, C_1, C_2) = U(1-L, v(C_1, C_2)). \tag{46}$$

These preferences are separable in leisure and consumption, so if we translate the analysis of Atkinson and Stiglitz (1976) to an intertemporal setting, interpreting C_1 and C_2 as two different commodities, we may conclude that these two goods should be uniformly taxed, given that the government can pursue its distributional goals via the nonlinear labour income tax. This was also the conclusion reached by Ordover and Phelps (1979) in an explicit overlapping generations setting, provided the government can use debt policy or public investment to steer the capital stock towards its Golden Rule level.

The policy relevance of this result is that, in the absence of firm knowledge about the degree of substitutability or complementarity between leisure and consumption in the two periods, a natural benchmark is to assume that present and future consumption are equally substitutable for leisure. As we have just seen, the marginal effective tax rate on capital income should then be set to zero. In practice this could be achieved by exempting capital income from tax, or by introducing a cash flow expenditure tax which is known to imply a zero marginal effective capital income tax rate (when the tax rate is constant over time).

However, as argued by Erosa and Gervais (2002), the leisure taken by the typical consumer does in fact tend to increase with age, suggesting that leisure and future consumption are complements. In that case the theory of optimal commodity taxation suggests that a positive capital income tax could be part of an optimal tax system, assuming that the life cycle model provides an adequate description of intertemporal consumer behaviour.

An alternative vision of consumer behaviour is embodied in the popular infinite horizon model originally suggested by Ramsey (1928) and usually defended by the assumption that generations are linked via an altruistic bequest motive so that consumers effectively behave as if they had an infinite horizon. Within this framework Chamley (1986) and Judd (1985) found that the optimal steady state tax rate on capital income is zero even if the alternative to capital income taxation is a distortionary labour income tax. One way of explaining this result is to note that in the Ramsey model the steady-state after-tax interest rate is closely tied to the consumer's exogenous utility discount rate. With a constant long-run equilibrium after-tax interest rate, a capital income tax gets fully shifted onto the pre-tax interest rate, i.e. the supply of capital is in effect infinitely elastic in the long run. Clearly it is not optimal to tax a factor with an infinite elasticity of supply. Further, regardless of the relevance of the Ramsey infinite horizon model, a small open economy with perfect capital mobility also faces an infinitely elastic supply of capital from the world capital market and hence should levy no source-based taxes on capital, as argued by Razin and Sadka (1991).

However, for a number of reasons discussed at length in Sørensen (2007a), the result that the optimal capital income tax rate is zero is not robust. For example, if pure profits cannot be fully taxed away, a source-based capital income tax may be a second-best means of taxing location-specific rents generated by domestic economic activity (see Stiglitz and Dasgupta (1971) and Huizinga and Nielsen (1997)). The source-based corporate income tax may also be a necessary backstop to the personal income tax, since otherwise income could be accumulated free of tax in the corporate sector. Further, political constraints on the use of other sources of revenue may simply force the government to rely to some extent on the revenue from capital income taxes.

Should capital income taxes be neutral?

It is usually argued that a capital income tax should be neutral in the sense that all returns to capital should face the same marginal effective tax rate. In other words, capital income taxes should not be differentiated across different sectors or economic activities. This may be seen as an application of the famous production efficiency theorem of Diamond and Mirrlees (1971) which says that the second-best optimal tax system avoids production distortions provided the government can tax away pure profits and can tax all transactions between households and firms.

In an open economy production efficiency requires that the government can tax all returns to capital received by domestic residents from foreign as well as domestic sources. But enforcing a residence-based tax on worldwide income requires an effective system of information exchange between national tax administrations, and such a system does not exist in practice, not even in the European Union.[13] Tax administrators therefore find it hard to monitor the taxpayers' foreign source capital income, so in practice capital income taxes tend to be source-based. The domestic capital income tax thus becomes a selective tax on domestic investment as opposed to investment abroad. This violates the assumptions underlying the Diamond-Mirrlees production efficiency theorem and hence there is no longer any presumption that all domestic activities should face the same marginal effective capital income tax rate. Instead, there is a case for differentiating the capital income tax in accordance with Ramsey principles.

[13] The EU Savings Tax Directive does require exchange of information on interest income received by EU personal taxpayers from other EU countries (and from a number of co-operating jurisdictions), but there are a number of ways in which taxpayers can in practice escape information exchange, e.g. by placing a legal person or a trust or fiduciary between the beneficial owner and the interest-generating asset. See the European Commission (2008) for details on this.

A Ramsey rule for optimal source-based capital income taxation

To see this, consider a small open economy where total output *Y* is given by the sum of the outputs Y_1 and Y_2 from two production sectors using capital inputs K_1 and K_2 plus a fixed factor.[14] The outputs are given by the concave production functions

$$Y_1 = f\left(K_1\right), \quad f' > 0, \quad f'' < 0, \qquad\qquad Y_2 = F\left(K_2\right), \quad F' > 0, \quad F'' < 0. \tag{47}$$

Capital is perfectly mobile across borders, and investors can earn a fixed after-tax return *r* by investing abroad. If t_i is the unit tax on capital invested in domestic sector *i*, the marginal pre-tax return on investment in that sector must therefore equal $r + t_i$ to ensure an after-tax return equal to that obtainable by investing abroad. Hence we have the capital market equilibrium conditions

$$f'\left(K_1\right) = r + t_1, \qquad\qquad F'\left(K_2\right) = r + t_2, \tag{48}$$

which imply that the capital stocks invested in the two sectors vary negatively with the cost of capital:

$$K_1 = K_1\left(r + t_1\right), \quad K_1' = \frac{1}{f''} < 0, \qquad\qquad K_2 = K_2\left(r + t_2\right), \quad K_2' = \frac{1}{F''} < 0. \tag{49}$$

Assuming a given national wealth endowment $\bar{K}$, the amount of domestically-owned capital invested abroad is $\bar{K} - K_1 - K_2$, so total national income is

$$\begin{aligned} Y &= Y_1 + Y_2 + r\left(\bar{K} - K_1 - K_2\right) \\ &= f\left(K_1\left(r + t_1\right)\right) + F\left(K_2\left(r + t_2\right)\right) + r\left[\bar{K} - K_1\left(r + t_1\right) - K_2\left(r + t_2\right)\right]. \end{aligned} \tag{50}$$

The optimal tax system maximises national income (50) with respect to t_1 and t_2 subject to the constraint that the government must raise a given amount of revenue *R*:

$$t_1 \cdot K_1\left(r + t_1\right) + t_2 \cdot K_2\left(r + t_2\right) = R. \tag{51}$$

The solution to this problem can be written in either of the following ways:

$$\varepsilon_1 = \varepsilon_2, \qquad\qquad \varepsilon_i \equiv \frac{dK_i}{dt_i}\frac{t_i}{K_i}, \qquad i = 1, 2, \tag{52}$$

$$\frac{t_i}{r + t_i} = \frac{\lambda}{\eta_i}, \qquad\qquad \eta_i \equiv -\frac{dK_i}{d\left(r + t_i\right)}\frac{r + t_i}{K_i}, \qquad i = 1, 2. \tag{53}$$

Equation (52) is a *Ramsey rule for optimal capital taxation* stating that, at the margin, the tax system should cause the same relative reduction of investment in the different production sectors. In

[14] The model is inspired by the seminal general equilibrium framework developed by Harberger (1962). A more general version of the model, incorporating three factors of production, is analysed in Sørensen (2007b); however, the qualitative conclusions remain the same as those reported below.

general, this rule calls for *differentiated* capital taxation since the elasticity of capital demand with respect to the tax rate (ε_i) will generally differ across sectors. The optimal degree of tax differentiation is given by the *inverse elasticity rule* (53) where λ is the shadow price of public funds. This rule says that the marginal effective capital income tax rate $t_i/(r+t_i)$ on a given sector should be inversely proportional to that sector's elasticity of capital demand with respect to the cost of capital (η_i). Obviously this policy rule is analogous to Ramsey's famous inverse elasticity rule for optimal commodity taxation.

If production functions take the constant-elasticity forms $Y_1 = AK_1^{\alpha}$, $0<\alpha<1$, and $Y_2 = BK_2^{\beta}$, $0<\beta<1$, it follows from (48) that $\eta_1 = 1/(1-\alpha)$ and $\eta_2 = 1/(1-\beta)$, implying that the optimal relative tax rates given by (53) become

$$\frac{t_1/(r+t_1)}{t_2/(r+t_2)} = \frac{1-\alpha}{1-\beta}. \tag{54}$$

In competitive markets the parameters α and β are equal to the capital income shares of total sectoral income which can be estimated from the national income accounts. We see that the more capital-intensive sector (measured by the capital income share) should carry a lower relative tax rate.

Optimal capital taxation and capital mobility

With the constant elasticity production functions assumed above, the solutions for capital stocks given in (49) take the form

$$K_1 = \left(\frac{r+t_1}{\alpha}\right)^{\frac{1}{\alpha-1}}, \qquad K_2 = \left(\frac{r+t_2}{\beta}\right)^{\frac{1}{\beta-1}}, \tag{55}$$

so that the tax elasticities in (52) become

$$\varepsilon_1 = -\left(\frac{1}{1-\alpha}\right)\left(\frac{t_1}{r+t_1}\right), \qquad \varepsilon_2 = -\left(\frac{1}{1-\beta}\right)\left(\frac{t_2}{r+t_2}\right). \tag{56}$$

From (55) it follows that a higher tax rate on one sector does not channel more capital into the other domestic sector. Hence the elasticities in (56) arise only from the international mobility of capital. We see that the greater the capital intensity of production (measured by the size of α and β), the larger is the capital outflow generated by a given tax rate, so the lower is the optimal sectoral tax rate.

This is consistent with the observation that governments often tend to offer relatively generous tax treatment of activities that are perceived to be particularly mobile across borders. For example, in practically all countries the shipping industry benefits from large tax concessions. Firms in this sector are indeed highly capital intensive and internationally mobile, in accordance with the analysis above.

It should be stressed once again that, in the model framework described here, the optimal policy for the world as a whole would be a system of automatic international information exchange enabling governments to enforce a uniform residence-based capital income tax that would preserve production efficiency. But in the absence of the worldwide co-operation needed to implement such a regime, governments are exposed to the forces of international tax competition unleashed by source-based taxation. From a national perspective standard Ramsey principles familiar from the theory of optimal commodity taxation may then provide a rationale for differentiated capital income taxation in order to minimise distortions to the domestic pattern of investment.

4. Concluding remarks

This paper has offered some examples from the recent literature on optimal taxation in the hope to convince the reader that this body of theory can in fact offer useful guidance for practical tax policy. The first part of the paper illustrated how an optimal non-linear labour income tax schedule may be derived from estimates of labour supply elasticities on the intensive and the extensive margin combined with data on the wage distribution plus assumptions on the relative marginal social evaluation of income for different income groups. In particular, we saw how recent advances in the theory of optimal labour income taxation could help governments to design optimal in-work benefits of the type that has become increasingly popular in OECD countries.

The second part of the paper dealt with the optimal design of indirect taxes. We noted that governments may often lack the information on cross-price elasticities with leisure needed to implement the differentiated commodity taxes prescribed by conventional optimal tax theory. This fact, combined with administrative and political economy considerations, suggests that uniform taxation should be the main guideline for indirect taxation. However, based on recent analyses of optimal taxation with household production, we argued that certain consumer services that can easily be produced as a do-it-yourself activity within the household should probably be subject to relatively low indirect tax rates in order to distort the pattern of market activities as little as possible.

The third part of the paper discussed the optimal taxation of income from capital, focusing on the issue whether capital income taxes should be “neutral”, i.e. uniform across all types of investment. We found that neutral taxation should indeed be the norm if international information exchange enables governments to enforce residence-based capital income taxation. However, absent an effective system of automatic information exchange, governments are typically forced to rely on source-based taxes on capital that discourage domestic investment as opposed to investment abroad. In that case the optimal system of capital income taxation secures an equal proportional reduction of investment in the different domestic production sectors. This principle generally calls for differentiated taxation in accordance with a Ramsey-type rule requiring that the marginal effective tax rate on capital income from a given sector be inversely proportional to that sector’s elasticity of capital demand with respect to the cost of capital. We also saw that the elasticity of capital demand reflects the degree to which domestic taxation induces a capital export, so the Ramsey rule for optimal capital income taxation provides a rationale for relatively low capital tax rates on activities that are highly mobile across borders. As in the case of indirect taxation, administrative and political economy concerns suggest that tax economists should be cautious when advising departures from uniform capital income taxation. For example, extensive differentiation of capital income tax rates would invite tax avoidance through transfer-pricing within groups of affiliated companies. The work of Auerbach (1989) also indicates that the efficiency gains from an optimal system of differentiated capital income taxes may generally be small. The analysis in this paper nevertheless suggests that in the case of highly mobile activities, deviations from “neutrality” may be warranted from a national perspective as long as national governments cannot effectively coordinate their tax policies.

References

Atkinson, A.B. & J. Stiglitz (1976). The design of tax structure: direct versus indirect taxation. Journal of Public Economics 6, 55-75.

Atkinson, A.B. & J. Stiglitz (1980). Lectures on Public Economics. New York: McGraw-Hill.

Auerbach, A.J. (1989). The deadweight loss from 'non-neutral' capital income taxation. Journal of Public Economics 40, 1-36.

Becker, G.S. (1965). A theory of the allocation of time. Economic Journal 75, 493-517.

Boadway, R. & F. Gahvari (2006). Optimal taxation with consumption time as a leisure or labor substitute. Journal of Public Economics 90, 1851-1878.

Boone, J. and A.L. Bovenberg (2004). The optimal taxation of unskilled labor with job search and social assistance. Journal of Public Economics 88, 2227-2258.

Browning, M. & C. Meghir (1991). The effects of male and female labour supply on commodity demands. Econometrica 51, 925-951.

Christiansen, V. (1984). Which commodity taxes should supplement the income tax? Journal of Public Economics 24, 195-220.

Chamley, C. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54, 607-622.

Corlett, W.J. & D.C. Hague (1953). Complementarity and the excess burden of taxation. Review of Economic Studies 21, 21-30.

Creedy, J. (2009). The personal income tax structure: theory and policy. Paper presented at the conference on New Zealand Tax Reform – Where to Next? at the Victoria University of Wellington, 11-13 February, 2009.

Diamond, P.A. (1975). A many-person Ramsey tax rule. Journal of Public Economics 4, 335-342.

Diamond, P.A. (1998). Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates. American Economic Review 88, 83-95.

Diamond, P.A. & J.A. Mirrlees (1971). Optimal taxation and public production (I). American Economic Review 61, 8-27.

Edgeworth, F.Y. (1897). The pure theory of taxation. Economic Journal 7, 550-571.

Erosa, A. & M. Gervais (2002). Optimal taxation in life-cycle economies. Journal of Economic Theory 105, 338-369.

European Commission (2008). Report from the Commission to the Council, COM (2008) 552 final.

Harberger, A.C. (1962). The incidence of the corporation tax. Journal of Political Economy 70, 215-240.

Heckman, J.J. (1993). What has been learned about labor supply in the past twenty years? American Economic Review (AEA Papers and Proceedings) 83, 116-121.

Hosios, A. (1990). On the efficiency of matching and related models of search and unemployment. Review of Economic Studies 57, 279-298.

Huizinga, H. & S.B. Nielsen (1997). Capital income and profits taxation with foreign ownership of firms. Journal of International Economics 42, 149-165.

Hungerbühler, M., E. Lehmann, A. Parmentier and B. van der Linden (2006). Optimal redistributive taxation in a search equilibrium model. Review of Economic Studies 73, 743-767.

Judd, K. (1985). Redistributive taxation in a simple perfect foresight model. Journal of Public Economics 28, 59-83.

Kaplow, L. (2006). On the undesirability of commodity taxation even when income taxation is not optimal. Journal of Public Economics 90, 1235-1250.

Kaplow, L. (2008). The Theory of Taxation and Public Economics. Princeton and Oxford: Princeton University Press.

Kleven, H.J. (2004). Optimum taxation and the allocation of time. Journal of Public Economics 88, 545-557.

Kleven, H.J., W. Richter and P.B. Sørensen (2000). Optimal taxation with household production. Oxford Economic Papers 52, 584-594.

Laroque, G. (2005). Indirect taxation is superfluous under separability and taste homogeneity: A simple proof. Economics Letters 87, 141-144.

McCulloch, J.R. (1845). A Treatise on the Principles and Practical Influence of Taxation and the Funding System. (Edited with introduction of D.P. O'Brien, 1975). London: Longman.

Mirrlees, J.A. (1971). An exploration in the theory of optimum income taxation. Review of Economic Studies 38, 175-208.

O'Brien, D.P. (2009). Introduction. In Taxation and the Promotion of Human Happiness: An Essay by George Warde Norman (Edited by D.P. O'Brien with J. Creedy). Cheltenham: Edward Elgar.

Ordover, J. and E. Phelps (1979). The concept of optimal taxation in the overlapping generations model of capital and wealth. Journal of Public Economics 12, 1-26.

Petersen, L.B. (2007). Velfærdsvægte for forskellige indkomstgrupper i Danmark. *Nationaløkonomisk Tidsskrift* (*Danish Journal of Economics*) 145, 265-287.

Piketty, T. (1997). La redistribution fiscale face au chômage. Revue Francaise d'Economie, 12, 157-201.

Ramsey, F. (1927). A contribution to the theory of taxation. Economic Journal 37, 47-61.

Ramsey, F. (1928). A mathematical theory of saving. Economic Journal, 38, 543-559.

Razin, A. and E. Sadka (1991). International tax competition and gains from tax harmonization. Economics Letters 37, 69-76.

Revesz, J.T. (1989). The optimal taxation of labour income. Public Finance 44, 453-475.

Roberts, K. (2000). A reconsideration of the optimal income tax. In Hammond, P.J. and Myles, G.D. (eds.), Incentives and Organization: Essays in Honour of Sir James Mirrlees. Oxford: Oxford University Press.

Sadka, E. (1977). A theorem on uniform commodity taxation. Journal of Public Economics 7, 387-391.

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

Saez, E. (2002a). Optimal income transfer programs: intensive versus extensive labour supply responses. The Quarterly Journal of Economics 117, 1039-1073.

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

Sandmo, A. (1974). A note on the structure of optimal taxation. American Economic Review 64, 701-706.

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

Sandmo, A. (1983). Progressive taxation, redistribution and labor supply. Scandinavian Journal of Economics 85, 311-323.

Sandmo, A. (1990). Tax distortions and household production. Oxford Economic Papers 42, 78-90.

Spadaro, A. (2008). Optimal taxation, social contract and the four worlds of welfare capitalism. Paris School of Economics, Working Paper No. 2008-38.

Stiglitz, J. & P. Dasgupta (1971). Differential taxation, public goods, and economic efficiency. Review of Economic Studies 38, 151-174.

Sørensen, P.B. (1997). Public finance solutions to the European unemployment problem? Economic Policy 25, 223-264.

Sørensen, P.B. (1999). Optimal tax progressivity in imperfect labour markets. Labour Economics 6, 435-452.

Sørensen, P.B. (2007a). Can capital income taxes survive? And should they? CESifo Economic Studies 53, 172-228.

Sørensen, P.B. (2007b). The theory of optimal taxation: what is the policy relevance? International Tax and Public Finance 14, 383-406.

Tuomala, M. (1984). On the optimal income taxation: some further numerical results. Journal of Public Economics 23, 351-366.