ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Blomquist, Sören; Christiansen, Vidar

Working Paper

Taxation and heterogeneous preferences

CESifo Working Paper, No. 1244

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Blomquist, Sören; Christiansen, Vidar (2004) : Taxation and heterogeneous preferences, CESifo Working Paper, No. 1244, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/18883

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TAXATION AND HETEROGENEOUS PREFERENCES

SÖREN BLOMQUIST
VIDAR CHRISTIANSEN

CESIFO WORKING PAPER NO. 1244
CATEGORY 1: PUBLIC FINANCE
JULY 2004

An electronic version of the paper may be downloaded
- *from the SSRN website: www.SSRN.com*
- *from the CESifo website: www.CESifo.de*

CESifo *Working Paper No. 1244*

TAXATION AND HETEROGENEOUS PREFERENCES

Abstract

Non-linear income taxes and linear commodity taxes are analysed when people differ with respect to ability, high-skilled agents have heterogeneous preferences, and neither individual abilities nor preferences are observable. The paper highlights how informational constraints may motivate differential treatment of people with different preferences for leisure even if unequal treatment is not desirable per se. Which preference type will be better or worse off is shown to depend on the self-selection constraints associated with the information asymmetry.

We characterize pure income tax optima, which may be bunching or separating optima. In particular, the income tax may not be able to distinguish between those low-income people who are low-skilled and those who have strong preference for leisure. As is shown, there may still be an impact on the optimum income tax schedule as it will depend on the composition of the population with respect to types of individuals. Finally, the paper addresses what can be achieved by commodity taxes when preferences are heterogeneous, in particular, in terms of targeting groups that the income tax is incapable of discriminating between.

JEL Code: H21, H23.

Keywords. optimum taxation, heterogeneous preferences, asymmetric information.

Sören Blomquist
Department of Economics
Uppsala University
Box 513
751 20 Uppsala
Sweden
Soren.Blomquist@nek.uu.se

Vidar Christiansen
Department of Economics
University of Oslo
P.O.Box 1095 Blindern
0317 Oslo
Norway
vidar.christiansen@econ.uio.no

Previous versions of the paper have been presented at the CESifo public sector area conference, a staff seminar at the University of Dortmund, the IIPF conference in Prague 2003 and the Conference on "Welfarist and non-welfarist approaches to public economics" in honour of Tony Atkinson at Ghent University in March 2004. We are grateful for comments from the seminar and conference participants, and in particular to Efraim Sadka, Wolfram Richter and Luca Micheletto.

1. Introduction

There can be little doubt that in practice people are endowed with heterogeneous preferences and therefore choose different consumption bundles even if they have the same income or wage rate and face the same prices. Some have a strong taste for travelling, dining at restaurants, and going to the theatre, whilst others prefer to invest in housing and household durables to enjoy more time at home. Some prefer to put a lot of effort into a highly paid job, whilst others prefer a quiet and less stressful life. Although it is easy to recognise the variety of tastes in the population, optimum tax theory has been surprisingly silent on this score. Whilst a standard assumption is that various agents have different endowments or skill levels that motivate a distributional role for taxes, it has almost invariably been assumed that the same utility function motivates individual choices. An important question is whether neglecting differences in preferences is justified or whether the recognition of heterogeneity of preferences should impact the choice of tax policy. Should the income tax schedule be responsive to the distribution of preferences? Does preference heterogeneity make a difference when it comes to commodity taxes?

To address these questions we shall consider a population of individuals differing along two dimensions. There are low-skilled and high-skilled individuals with the latter group being comprised of people with strong preference for leisure and people with weak preference for leisure. All low-skilled are assumed to have weak preference for leisure.

Roughly speaking there are two approaches to the issue of how to treat equally skilful people with different preferences. One is what we may call the horizontal equity approach, which imposes the restriction that equally skilful individuals should be treated in some sense equally, for instance by being faced with the same tax liability irrespective of preferences. The other one is the social welfare approach, which implies that a social welfare function for all individuals, with their respective skills and preferences, is maximised even if that may imply different treatment of people with identical skill levels. The discussion of how to treat different types of people is often taking place under the explicit or implicit assumption that there is full information about individual skills and preferences. It is then conceivable that one may impose a high, and possibly uniform, tax on high-skilled individuals irrespective of preferences, and tax leniently, or make a transfer to, low-skilled individuals.

When there is asymmetric information about individual skills and preferences the perspective may be different. Even if at a no-tax point of departure there is considered to be no case for redistribution between equally skilful people, the case for equal treatment may be weakened as soon as vertical redistribution takes place, and asymmetric information constrains how taxes can be imposed. Assume that one starts redistributing from high-skilled people to low-skilled people. In the extreme case high-skilled people with strong preference for leisure and low-skilled individuals will make the same choice of observable income points. Any income point intended for the low-skilled will be adopted by the high-skilled with strong preference for leisure, and the equal treatment of all high-skilled breaks down. An attempt to redistribute from all high-skilled to low-skilled is not feasible as the high-skilled with strong preference for leisure will in fact be treated as low-skilled. One cannot at the same time treat all high-skilled equally and high-skilled and low-skilled differently. Vertical redistribution and horizontal equity becomes irreconcilable objectives.

In less extreme cases different income points may be assigned to the high-skilled with strong preferences for leisure and low-skilled individuals, respectively, but only subject to a self-selection constraint. Attempted redistribution from high-skilled to low-skilled may induce high-skilled with strong preference for leisure to start mimicking the low-skilled. The income point assigned to the low-skilled must be sufficiently distorted to discourage the high-skilled with strong preference for leisure from mimicking by choosing the same income point. The implication is that the transfers to the low-skilled are costly in terms of distortions. However, the cost can be reduced by lowering the transfer from the high-skilled with strong preference for leisure, who is the potential mimicker, whilst relying more on transfers from the high-skilled with weak preference for leisure. This will alleviate the cost of transfers to the low-skilled but only by abandoning equal treatment of the high-skilled with different preferences (horizontal equity). The general argument is that transfers from different types of high-skilled individuals may imply different costs under asymmetric information and hence employing the less costly alternative may be preferable, but only by violating horizontal equity.

Our conclusion is that it may be crucial to distinguish between full information and asymmetric information regimes when it comes to discussing the tax treatment of equally skilful individuals. It may be that horizontal equity can be enforced without serious problems under full information, or there may be no social welfare case for

redistribution between different high-skilled persons under full information, but this may be entirely different under asymmetric information. Imposing horizontal equity may then be costly. A social welfare approach may treat the various types of high-skilled individuals differently, not because that is desirable *per se*, but because it is part of a scheme involving transfers to low-skilled individuals. Assuming that we do indeed treat the high-skilled individuals differently by taxing more leniently those with strong preference for leisure (in order to reduce the cost of vertical redistribution), it is plausible that different social marginal utilities of income will be assigned to the high-skilled individuals even if no such discrepancy would exist in the no-tax regime. Below we shall pursue the welfare approach to elaborate on conceivable income tax optima when high-skilled individuals may be taxed differently.

Heterogeneity of preferences can, roughly classified, be of two kinds. People can have different trade-offs between income and leisure, or they can have different preferences for consumption bundles bought in the market, in the sense that they will buy different consumption bundles even if they have the same income and leisure and face the same prices. We may note that various combinations of the two types of preferences are conceivable. People may be endowed with different preferences for leisure, while still having the same preferences for market goods. People may also have different preferences for leisure *because* they have different preferences for market goods. The reason is simple. Since income is used for buying market goods, the preferences for market goods may affect the trade-off between income and leisure. We shall survey preference heterogeneity in further detail below.

As has been discussed by some authors (e.g. Cuff (2000)), heterogeneity of preferences can be given different interpretations. A stronger preference for a good may sometimes be interpreted as meaning that a person has a stronger need for the good in question. Indeed the person may have a disadvantage or handicap that can only be alleviated, at least partially, by consuming some good. In such a case we will perceive this person as being worse off if he does not get more than others of the good he has a special need for, and even then he may not be fully compensated for his handicap. A clear-cut example may be a person in need of medication. On the other hand it is also conceivable that a sort of handicap may induce a weaker preference for a commodity. An illiterate person will have a weak preference for books, as opposed

to a person with an intellectual mind. Thus a strong preference for a good may be interpreted either as reflecting a handicap or a particular ability to enjoy a good.

We will consider heterogeneity of preferences that neither reflects a superior nor an inferior overall endowment. People are simply different, without being systematically more or less well endowed physically, mentally or emotionally. They have different pure tastes. A central assumption is that the government cannot observe an individual's preferences. The government knows the distribution of preferences, but not the identity of people with the respective tastes.

The very small literature on taxation of a population of agents with heterogenous preferences has adopted various approaches. Sandmo (1993) asked how one should tax people who are endowed with the same resources, but have different preferences. Even in this very simplified setting there are few general conclusions to be drawn. Tarkiainen and Tuomala (1999) used numerical specifications to derive computationally optimum non-linear tax schedules for a continuum of agents simultaneously distributed by skill level and preferences for income and leisure. Their findings are interesting as illustrations of possible optima. However, their results could be sensitive to the specifications that have been made[2]. Assuming a special class of utility functions, Boadway et al. (2002) considered a discrete distribution of individuals asssuming there are two skill levels and two types of preferences. However the high-skilled person with strong preference for leisure cannot be distinguished from the low-skilled person with weak preference for leisure and must always be treated equally. The aim of their analysis is to trace the income tax implications of assigning different welfare weights to the high-skilled person with strong preference for leisure.

Saez (2002) discusses the role of commodity taxes supplementing an income tax in a very general framework with a continuum of individuals from a multidimensional set. The focus of his study is to identify conditions under which the Atkinson-Stiglitz (1976) result of zero commodity taxes can be recovered in a setting with heterogeneous preferences. The present paper is rather concerned with cases in which Saez' generalised Atkinson-Stiglitz conditions are not fulfilled, and accordingly there is scope for discussing the need for, and characterizing, the optimality conditions for commodity taxes. In contrast to Saez the present paper uses

[2] A main finding is that preference heterogeneity does not necessarily lower the progressivity of the income tax.

a model with a small, discrete number of types of individuals. Jordahl and Micheletto (2002) discussed tax policy implications of imposing a horizontal equity requirement when preferences for leisure vary. Cuff (2000) considered heterogeneous preferences in the context of workfare adopting a special class of utility functions.

In view of the existing literature, and in particular the discussion of Sandmo (1993) an appropriate question seems to be: Why is it so hard to draw conclusions about distributional policy when preferences vary? Part of the answer seems to be that making interpersonal utility comparisons is particularly hard when preferences are acknowledged to be heterogeneous[3]. Even if in principle we find ourselves able to compare utilities, who is judged to be better off may depend on the parameters of the economy such as prices and wage rates. Even if person *A* is judged to be better off than person *B* for one set of parameters the difference may be evened out or reversed if there is a change to a different set of parameters. The implication is that even if we are willing to base utility comparisons on strong assumptions, there is the further difficulty of obtaining sufficiently robust assumptions, i.e. assumptions that are sufficiently invariant to the specific parameters of the economy we study. However, there is even a further challenge. As emphasised by Sandmo (1993) the direction of redistribution should be governed by differences in *marginal* rather than total utilities. This would be a rather uninteresting statement if there were a unique relationship between total and marginal utilities, but this will only be true under very restrictive assumptions. In general the marginal utility of income does not only depend on which indifference curve that is assigned to the person but also on the particular commodity bundle that generates the corresponding utility.

Throughout our discussion we shall make the assumption that, based on pure (marginal) utility comparisons it would be desirable to make further transfers to the unskilled individuals. The actual transfer is, however, limited by the informational constraints facing the decision-maker. This is in line with the standard assumption in the (conventional) tax theory with individuals of different skills, but homogeneous preferences. In our context we shall even assume that it is desirable to improve the welfare of the low-skilled when comparing with both high-skilled types. This implies a comparison across preferences and is a stronger assumption than the conventional one. However, we shall leave open the outcome of (marginal) utility comparisons

[3] This formidable task seems to require further empirical research in the border area of psychology and economics.

between the high-skilled individuals with different preferences, and consider alternative cases.

Heterogeneity of preferences can affect the choice of income tax schedule, the case for employing commodity taxes and the rates of these taxes. We shall address all these aspects.

In the next section we provide a taxonomy of preference heterogeneity. Section 3 addresses various pure income tax optima when preferences are heterogeneous, and examines the effects of changing the composition of the population. Section 4 discusses what commodity taxes can achieve when preferences are heterogeneous. Section 5 concludes.

2. A taxonomy of preference heterogeneity

The simplest models of labour supply assume that there is only a single market commodity, and preferences solely reflect the trade-off between the consumption of this homogeneous commodity and labour (or leisure). In general people have preferences over bundles of market commodities and labour, and these preferences may be identical, or heterogeneous in various ways. We may in general write the utility function as $u^i(z,h)$ where the superscript i indicates a particular type of preference, z denotes a vector of market commodities, and h is hours of work. Since there is a unique relationship between labour and leisure, we may choose to express preferences in terms of either of these variables. We opt for labour as argument in the utility function.

A problem with this general approach is that it covers a multitude of preferences, and without further restrictions almost any behavioural variation and any welfare considerations are possible. Therefore it is helpful to start by identifying certain classes of preferences, distinguishing between preferences for market goods on the one hand and the trade-off between market goods and leisure on the other hand.

Let us consider the following categories of preferences:

i. There are identical preferences for market commodities, but different preferences for labour, which is assumed to be weakly separable from market goods in the utility function. The utility function belongs to the class $u^i(z,h) = u^i(f(z),h)$. Choosing his bundle of market commodities, the consumer will maximise $f(z)$ s.t. the budget constraint $B=pz$, with B denoting after tax income and p the price vector, and we get

the maximum value function *f*(p, B)*. The preferences for after tax income and labour are then expressed by $v^i(p,B,h) = u^i(f^*(p,B),h)$.

ii) Labour is not weakly separable from market goods, but conditional on labor there are uniform preferences for market goods, while preferences for labour differ. The utility function is of the form $u^i(z,h) = u^i(f(z,h),h)$, assumed to be increasing in the first argument and deceasing in the second. The preferences for after tax income and labour are expressed by $v^i(p,B,h) = u^i(f^*(p,B,h),h))$.

iii) Labour is weakly separable from market commodities. Preferences for market goods are heterogeneous, whilst people have the same preferences for leisure and market commodity aggregates. The utility function is of the form $u^i(z,h) = u(f^i(z),h)$, where the interpretation is that it is the ordinal properties of the *f*-function that may vary between consumers[4]. The preferences for after tax income and labour are expressed by $v^i(p,B,h) = u(f^i{}^*(p,B),h)$.

iv) Labour is not weakly separable from market commodities, and preferences for market commodities are heterogeneous. Obviously no separability is obtained in the corresponding v^i-function.

On the basis of this discussion we recognise that the income/labour trade-off expressed by a utility function *v(B,h)* can originate from several underlying basic preference patterns from which its properties are derived. In particular, we note how (non-)separability between after tax income and labour depends on the properties of the underlying basic utility function.

Most previous contributions have abstracted from the plurality of market commodities and have just addressed preferences for disposable income and labour without discussing how the disposable income is spent. We may note that when there are underlying preferences for multiple commodities the corresponding price vector affects the location and shape of indifference curves in *h,B*-space.

[4] A specific example might be $z_1^{a_i} z_2^{1-a_i} - h^b$ where a_i and b are positive parameters and a_i is type specific.

3. Income taxation when there is heterogeneity in preferences between leisure and consumption

We will apply a three-type version of a Mirrlees (1971) optimal income tax model, which is similar to the extensively employed two-type model first introduced by Stern (1982) and Stiglitz (1982). Each type of person has a skill level, which is reflected by his exogenous wage rate, *w*. We will index the three types of individuals of our model economy in the following way: type *1* is a low-skilled person with weak preference for leisure, type *2* is a high-skilled person with strong preference for leisure and type *3* is a high-skilled person with weak preference for leisure. We will denote the wage rate of a low skill person by w^1 and that of a high skill persons by $w^2 = w^3$. A person of type *i* (=1,2,3) will earn a before-tax income $Y^i = w^i h^i$. Even if these statistical properties of the population are public information, we shall assume that there is asymmetric information in the sense that the government (tax authority) is informed neither about individual skill levels nor individual preferences.

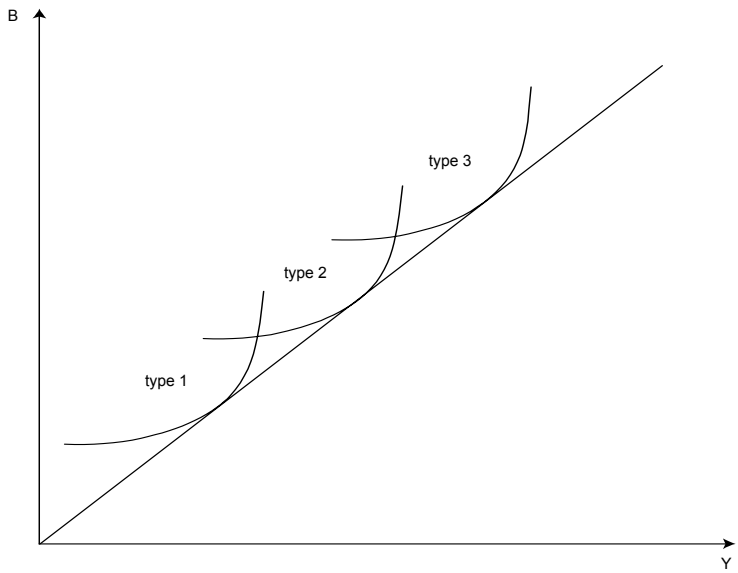

Figure 1

To give a first basic intuition for the labour-leisure trade-offs, we make use of figures 1 and 2. In figure 1 we depict a situation with no taxation. The assumptions implied by figure 1 is that at any point in the (Y, B) space type 1 has steeper indifference curves than type 2, and type 2 has steeper curves than type 3. Since the type 2 person has the same skill (wage rate) as the type 3 person, but works less and earns less

income, we say that the type 2 person has a strong preference for leisure. There are of course other possible patterns than the one illustrated in figure 1. For example, the type 1 person could have his earnings point between the earnings points of type 2 and type 3. However, the general insights do not depend on the exact pattern, so let us rather elaborate on the case illustrated in figure l.

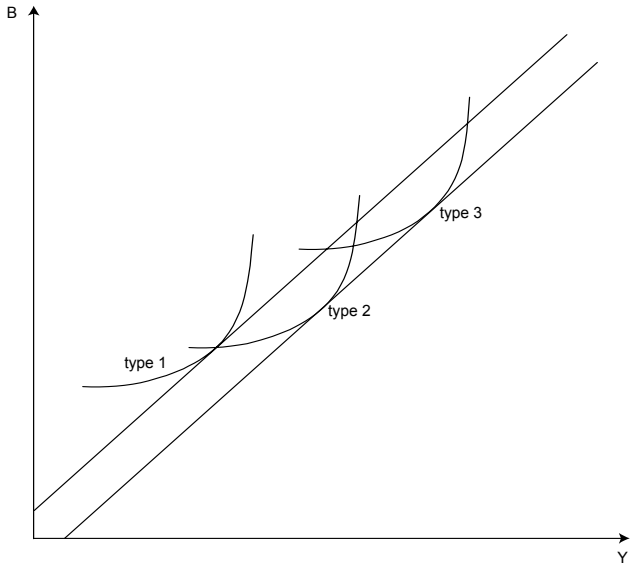

Figure 2

Our ethical premise is that redistribution from the high skill persons to the low skill persons is warranted. However, as was the conclusion of Sandmo (1993) we cannot in general tell whether we would like to redistribute between equally skilful individuals. In the absence of such knowledge, we shall assume there is indeed no reason to redistribute from one high skill person to the other unless the cost of redistribution to the low skill person differs between the two high skill types. If we are in the situation illustrated in figure 1 we can start taxing the type 2 and 3 persons by the same amounts, transferring the proceeds to the type 1 persons. This can go on in a non-distortionary way until we reach the situation illustrated in figure 2. Here the self-selection constraint that the type 2 persons should not mimic the type 1 persons is on the verge of binding. If we would like to push redistribution further, the cost will differ depending on whether we tax the type 3 or the type 2. The ethical premise is that in the situation of figure 2 the utility loss would be the same whether we take one dollar from type two or type three. In an appendix we argue that it is possible to establish a social welfare function that embodies this ethical assumption. However, if

we take a dollar from type 2 we can, due to the binding self-selection constraint, increase utility less for the type 1 person than if we take the dollar from the type 3 person. If we take another dollar from type 3 and transfer it to type 1, we can move out along the indifference curve of the type 2 person. If we take a dollar from the type 2 person, the indifference curve of type 2, shifts downwards, and it is along this lower level curve we can move the earnings point for the type 1 person. This implies a lower utility level for the type 1 person.

Hence, when we pursue further redistribution from the situation in figure 1, we will tax the type 3 more heavily than type 2, and there will be horizontal inequity. This is not because we want to redistribute between type 2 and type 3 *per se*, but because redistribution from type 3 to type 1 is more efficient.

Figures 1 and 2 illustrate one possible pattern of indifference curves. There can be other patterns and other self-selection constraints that bind. For example, the optimum could be such that the binding self-selection constraint is that the type 3 should not mimic type 1. In general, there can be two classes of optima, a screening optimum, where the three types of individuals are assigned three distinct income points, or a bunching optimum such that the type *1* and type *2* persons are assigned the same income point. Below we characterize a screening optimum in section 3.1 and a bunching case in section 3.2. We assume that the usual single crossing properties hold.

3.1 Screening optima

We will initially assume that indifference curves are such that the type *1* persons have the steepest slope as in the figures above. This implies that the type *2* person may mimic type *1* by choosing the income point intended for the latter, and conceivably type *3* may mimic type *2*. There are n^i persons of type *i*. The optimisation problem to be addressed is the maximisation of an additive welfare function (sum of utilities) subject to the budget and self-selection constraints. The corresponding Lagrange expression is

$$\Lambda = n^1 V^1(B^1,Y^1) + n^2 V^2(B^2,Y^2) + n^3 V^3(B^3,Y^3) + \alpha^3\left(V^3(B^3,Y^3) - V^3(B^2,Y^2)\right)$$
$$+\alpha^2\left(V^2(B^2,Y^2) - V^2(B^1,Y^1)\right) + \mu\left(n^1(Y^1 - B^1) + n^2(Y^2 - B^2) + n^3((Y^3 - B^3)\right) \quad (1)$$

where the Lagrange multiplier α^3 has been assigned to the self-selection constraint that type 3 should not mimic type 2, and α^2 is assigned to the self-selection constraint

that the high-skilled type with strong preference for leisure should not be mimicking the low-skilled. The last restriction is the government budget constraint.

The first order conditions are

$$B^1: \quad n^1 V_B^1 - \alpha^2 V_B^{21} - \mu n^1 = 0 \tag{2}$$

$$Y^1: \quad n^1 V_Y^1 - \alpha^2 V_Y^{21} + \mu n^1 = 0 \tag{3}$$

$$B^2: \quad n^2 V_B^2 - \alpha^3 V_B^{32} + \alpha^2 V_B^2 - \mu n^2 = 0 \tag{4}$$

$$Y^2: \quad n^2 V_Y^2 - \alpha^3 V_Y^{32} + \alpha^2 V_Y^2 + \mu n^2 = 0 \tag{5}$$

$$B^3: \quad n^3 V_B^3 + \alpha^3 V_B^3 - \mu n^3 = 0 \tag{6}$$

$$Y^3: \quad n^3 V_Y^3 + \alpha^3 V_Y^3 + \mu n^3 = 0 \tag{7}$$

Double superscripts indicate that a person, indicated by the former superscript, mimics some other person, indicated by the latter superscript. We denote the income tax function by *T(Y)*. It follows from eqs. (6) and (7) that the type three persons are undistorted facing a zero marginal tax $(T'(Y^3) = 0)$. For a low skill person we obtain the usual result of a positive marginal tax. Manipulations of eqs. (2) and (3) yield $T'(Y^1) = \rho_1 (MRS^1 - MRS^{21}) > 0$, where the marginal rate of substitution $MRS = -V_Y / V_B$, and ρ_1 denotes a positive term[5]. For the high skill type with a strong preference for leisure, the type 2 persons, manipulations of eqs. (4) and (5) yield a similar expression $T'(Y^2) = \rho_2 (MRS^2 - MRS^{32}) > 0$. In general we cannot tell which type has the higher marginal utility of income or the higher marginal income tax.

Let us now consider the special case in which the constraint that type three should not mimic type two is non-binding, whilst the other constraint remains effective. This implies that $\alpha_2 > 0$ and $\alpha_3 = 0$. The first order conditions imply

$$-V_Y^2 / V_B^2 = -V_Y^3 / V_B^3 = 1 \tag{8}$$

[5] We adopt the standard approach of equating the marginal rate of substitution to the marginal income net of tax, i.e. $-V_Y / V_B = 1 - T'(Y)$ by which the marginal tax rate is implicitly defined.

and

$$-\frac{V_Y^1}{V_B^1}-\frac{V_Y^{21}}{V_B^{21}}=\left(1-\frac{V_Y^1}{V_B^1}\right)\frac{\mu n^1}{\alpha^2 V_B^{21}}>0 \tag{9}$$

It is then easy to recognize that both high-skilled types will be undistorted, while the low-skilled type gets a distorted consumption bundle.

Simple manipulations of the conditions above then imply that

$$V_B^3=\mu\,, \tag{10}$$

$$V_B^1=\mu+\frac{\alpha^2 V_B^{21}}{n^1}>V_B^3\,, \tag{11}$$

$$V_B^2=\mu\frac{n^2}{n^2+\alpha^2}=V_B^3\frac{n^2}{n^2+\alpha^2}<V_B^3\,, \tag{12}$$

and accordingly

$$V_B^1>V_B^3>V_B^2\,. \tag{13}$$

Analogously we find when considering redistribution in terms of leisure (labour).

$$-V_Y^3=\mu\,, \tag{14}$$

$$-V_Y^1=\mu+\frac{-\alpha^2 V_Y^{21}}{n^1}>-V_Y^3\,, \tag{15}$$

$$-V_Y^2=\mu\frac{n^2}{n^2+\alpha^2}<-V_Y^3\,, \tag{16}$$

and accordingly

$$-V_Y^1>-V_Y^3>-V_Y^2\,. \tag{17}$$

We have assumed that transfers for the benefit of the low-skilled type are desirable to such an extent that he may be mimicked by high-skilled individuals. If only the mimicking constraint on the high-skilled type with strong preference for leisure is binding, the optimum is characterised by the social marginal utility of income being highest for the low-skilled and lowest for the high-skilled with strong preference for leisure.

The following intuition supports our understanding of this case. As one redistributes from high-skilled to low-skilled, at some point type two would like to mimic type one, and further transfers from *two* to *one* become more costly as it requires a distortion of type one's behaviour. This is a case for giving priority to transfers from type three, who is not (yet) induced to mimic. The larger transfers from

type three implies that at some point he is considered to be over-taxed compared to type two in the sense that ideally one would like to redistribute from type 2 to type 3. The reason why this imbalance persists at the optimum is that, if one were to tax type 2 harder, he would escape by mimicking type 1. Then there would be two options. *i.* One would have to worsen the situation also for type 1, which implies that a pure transfer from type 2 to type 3 is not feasible, and it is this additional cost in terms of lower welfare for the low-skilled that prevents the redistribution from taking place. *ii.* Alternatively one could maintain the utility of type 1, but only at the cost of further distorting the allocation, which would imply achieving the same utility at a higher resource cost. There would be a 'leakage' implying that for each income unit foregone by type 2, less than one unit would accrue to type 3.

Ideally one would also like to make further transfers from the high-skilled to the low-skilled. A further transfer from the high-skilled with stronger preference for leisure is blocked by the standard mimicking problem. Appropriating a unit of income from the high-skilled does not allow a one unit transfer to the low-skilled as a further distortion is required in order to rule out mimicking. Similarly a transfer from type 3 to type 1 would make the latter a more attractive target for mimicking by type 2.

We next focus on the welfare maximising case corresponding to (1) where the only self-selection constraint is that the industrious high skill person should not mimic the low-skilled. We let β^3 denote the Lagrange multiplier for this constraint. Analogous to the previous case, the first order conditions now imply

$$-V_Y^2 / V_B^2 = -V_Y^3 / V_B^3 = 1 \tag{18}$$

and

$$-\frac{V_Y^1}{V_B^1} - \frac{-V_Y^{31}}{V_B^{31}} = \left(1 - \frac{-V_Y^1}{V_B^1}\right)\frac{\mu n^1}{\beta^3 V_B^{31}} > 0 \tag{19}$$

where the sign follows from the standard assumption that the high-skilled type has the steeper indifference curve through any given point in *Y,B*-space. We get the conventional situation that the low-skilled person is distorted by a positive marginal tax rate at the optimum (efforts to increase gross income by one unit are rewarded by a smaller amount). Assuming no mimicking between high-skilled individuals, both high-skilled types are undistorted.

Now turning to distributional concerns, we note the further implications

$$V_B^1 - V_B^3 = \frac{\beta^3 V_B^{31}}{n^1} + \frac{\beta^3 V_B^3}{n^3} > 0 \tag{20}$$

$$V_B^1 = V_B^2 + \frac{\beta^3 V_B^{31}}{n^1} > V_B^2 \tag{21}$$

$$V_B^2 = \mu = \frac{n^3 + \beta^3}{n^3} V_B^3 > V_B^3 \tag{22}$$

Then $V_B^1 > V_B^2 > V_B^3$, (23)

And similarly

$-V_Y^1 > -V_Y^2 > -V_Y^3$. (24)

When only the mimicking constraint on the high-skilled type with weak preference for leisure is binding (i.e. type three is on the verge of mimicking), the optimum is characterised by the social marginal utility of income being highest for the low-skilled and lowest for the high-skilled with strong preference for leisure. In this case the difference of marginal utilities implies that ideally one would like to redistribute from type 3 to type 2. Once again it is the mimicking constraint that prevents such a transfer as it could not take place without adversely affecting the low-skilled person or adding to the resources required to ensure him the same utility as before. Also transfers from high-skilled to low-skilled would to some extent be eroded by mounting self-selection problems. For instance a transfer from type 2 to type 1 would require also 'bribing' type 3 or further distorting the transfer to person 1 to avert mimicking.

Our intuition for how this regime comes about is that, as income is transferred from high-skilled to low-skilled, at some point type three is induced to mimic type one, whilst type two is not. The latter has such a strong preference for leisure that he will choose an income which is low, and an amount of leisure which is large even when comparing with the low skilled. Redistribution to the low-skilled is then cheaper when transfers are made from type two rather than from type three since the latter will require a distortion of the behaviour of the low-skilled. This may shift the emphasis to transfers from type two implying an income distribution between two and three favouring type three.

3.2 Comparison of utility levels

It is often claimed that redistributive taxes will favour high-skilled people with a strong preference for leisure. A famous example is Nozick's question why someone preferring to look at the sunset should pay less tax than somebody who has to earn money in order to pay for material pleasures. (Nozick (1974)). It may therefore be of interest to compare utility levels at the social optimum. In both cases of a single binding self-selection constraint we have obtained an ordering of marginal utilities. However, in general no ordering of utility levels is available since the marginal utility of income does not only depend on the utility level, but also on the bundle of goods that generates the utility level. For a fixed utility level V, and for convenience measuring labour supply by gross income Y, we can express B by an expenditure function $e(Y,V)$. The slope of the indifference curve in Y,B-space is $dB/dY = e_Y$. Inserting $V(B,Y)$ we can write $B=e(Y,V(B,Y))$. Differentiating partially with respect to B we find the marginal utility of income $V_B = 1/e_V$. To trace the variation along the indifference curve we consider $\frac{\partial}{\partial Y}\frac{1}{e_V} = \frac{-e_{YV}}{e_V^2}$. It is easy to recognise that $e_{YV}>0$ if leisure is strictly non-inferior. The marginal utility of income will in general depend on the consumption-labour bundle, and it will decrease along an indifference curve as the labour supply and income increase if leisure is a strictly non-inferior good. It follows that a low marginal utility of income may be assigned to a person although he has a low utility level if he is sufficiently far out along his indifference curve.

The marginal utility of income will depend solely on the utility level (be constant along an indifference curve) only if there is no income effect on leisure

(e_{YV}=0). That means that the utility function belongs to the class $V=\psi(B\text{-}g(Y))$, and we see that $V_B = \psi'(B - g(Y)) = \psi'(\psi^{-1}(V)$. We may call this utility function B-quasi linear.

Two further assumptions will then allow us to draw conclusions about utility levels. Firstly we assume inequality aversion in the following sense.

Assumption I

For a given B-quasi-linear utility function the social marginal utility of income is decreasing in utility.

Secondly we assume the following relationship between preferences and marginal utility.

Assumption II

When different persons with B-quasi-linear preferences are considered to enjoy the same utility level, an individual with stronger preference for income derives more utility from additional income.

We can then easily establish the following insight.

Lemma 1: Given assumptions I and II, if preferences are B-quasi-linear, and the type of person with weaker preference for income has the higher marginal utility of income, he has the lower utility.

This is straightforward. If utility levels were the same, the person with stronger preference for leisure would have a lower marginal utility of income according to Assumption II. Since by Assumption I the marginal utility of income is decreasing in utility, it is only if the individual with stronger preference for leisure has a sufficiently lower utility that he will have the higher marginal utility of income.

Given the special assumptions above, it is then straightforward to conclude that when the only binding self-selection constraint is that the high-skilled individuals with stronger preference for leisure (type 2) do not mimic the low-skilled individuals (type 1), the high-skilled individuals with stronger preference for income have a lower utility than the high-skilled with stronger preference for leisure.

Analogously, it can be shown that the marginal utility of leisure is constant along an indifference curve only if the utility function is h-quasi-linear, i.e. it belongs to the class $V = \varphi(\xi(B) - h)$, where φ and ξ are monotonic, increasing functions[6].

Analogous to assumptions I and II we can state:

Assumption III

For a given h-quasi-linear utility function the social marginal utility of leisure is decreasing in utility.

Assumption IV

When different persons with h-quasi-linear preferences are considered to enjoy the same utility level, an individual with stronger preference for leisure derives more utility from additional leisure.

As an analogy to Lemma 1 we can then easily establish

Lemma 2: If preferences are h-quasi-linear, and the type of person with weaker preference for leisure has the higher marginal utility of leisure, he has the lower utility.

We can then conclude that under these special assumptions, when the only binding self-selection constraint is that the high-skilled individuals with stronger preference for income (type 3) do not mimic the low-skilled persons (type 1), the high-skilled with stronger preference for leisure enjoy lower utility than the high-skilled with stronger preference for income.

Even if the two cases considered are polar cases, that we may not find realistic, they serve to illustrate that in general we cannot tell whether it is the hardworking type or the type with a liking for leisure who is better off at the optimum. The outcome will depend both on the utility function and on which self-selection constraints that will bind.

---

[6] Let $Y = \varepsilon(B,V)$ be the labour (gross income) that in conjunction with a disposable income *B* will yield the utility level *V*. ε_B is the inverse of the slope of the indifference curve in *Y,B*-space. Differentiation of $Y = \varepsilon(B, V(B,Y))$ implies $-V_Y = -1/\varepsilon_V$, and $\frac{\partial}{\partial B}\frac{-1}{\varepsilon_V} = \frac{\varepsilon_{BV}}{\varepsilon_V^2}$, which is zero only if ε_{BV}=0.

The following propositions summarise the characterisations of the two regimes with a single binding self-selection constraint.

Proposition 1: When the only binding self-selection constraint is that the high-skilled individuals with stronger preference for leisure (type 2) do not mimic the low-skilled individuals (type 1),

a. both high-skilled individuals are undistorted while the low-skilled face a positive marginal tax rate,
b. the low-skilled individuals have the highest social marginal utility of income, and the high-skilled individuals with stronger preference for leisure have the lowest marginal utility of income.
c. assuming B-quasi-linear preferences and making assumptions *I* and *II*, the high-skilled individuals with stronger preference for income have a lower utility than the high-skilled with stronger preference for leisure.

Proposition 2: Let the only binding self-selection constraint be that the high-skilled individuals with stronger preference for income (type 3) do not mimic the low-skilled persons (type 1). Then

a. both high-skilled individuals are undistorted while the low-skilled face a positive marginal tax rate,
b. the low-skilled individuals have the highest social marginal utility of income, and the high-skilled individuals with stronger preference for income have the lowest social marginal utility of income,
c. assuming h-quasi linear preferences and making assumptions II and IV, the high-skilled with stronger preference for leisure enjoy a lower utility than the high-skilled with stronger preference for income.

3.3 The bunching case and comparative statics.

High-skilled people with strong preference for leisure may be hard to distinguish from low- skilled people. As a polar case we may consider the situation in which type *1* and type *2* have identical indifference maps in *Y,B*-space.[7]

[7] A special case would be the one presented by Boadway et al. (2002).

Let Y^i, B^i be the (gross and net) income bundle of type *i*. Since type *1* and type *2* have identical indifference curves their income points must be on the same indifference curve. Otherwise, one of the income points would be strictly preferred by both types, and the other point could not be imposed through self-selection. Moreover, the same income on the common indifference curve must be assigned to both types at the optimum. If not, there would always be a gain from switching one of the individuals to the less resource consuming of the two points. Hence we can set $Y^2 = Y^1$ and $B^2 = B^1$. Since an income tax cannot discriminate between types *1* and *2*, the only interesting tax issue is how the existence of different preferences will affect the respective taxes on category 3 on the one hand and on categories *1* and *2* on the other.

We shall confine our attention to the information-constrained optima where the choice of policy is strictly constrained by the requirement that the high-skilled type with weak preference for leisure prefers the income point Y^3, B^3 actually intended for him.

It is helpful to introduce the expenditure function $e^i(Y,V)$ expressing the disposable income required by a person of type *i* in order to achieve the utility level *V* when his gross income is *Y*. We can describe the economy by the following set of equations.

$$n^3(Y^3 - B^3) + m(Y^1 - B^1) = 0 \tag{25}$$

$$B^1 = e^1(Y^1, V^1) \tag{26}$$

$$B^3 = e^3(Y^3, V^3) \tag{27}$$

$$B^1 = e^3(Y^1, V^3) \tag{28}$$

(25) is the budget constraint with *m* $(= n^1 + n^2)$ being the number of low-income people (of type one and type two). (26) and (27) follow from the definition of the expenditure function, and (28) is the strictly binding self-selection constraint that the high-skilled person with weak preference for leisure obtain no higher utility when mimicking (choosing Y^1, B^1) than the level V^3 achieved when choosing the income point actually intended for him. We approach the social optimisation in two steps, first considering the Pareto optima and then the welfare maximisation Assigning any indifference curve (defined by V^1) to type *1* and *2*, Pareto optimality is achieved by

maximising the utility of type *3* s.t. the budget and self-selection constraint[8] and also pegging the value of V^1. We may note that the common indifference curve of type *1* and type *2* is steeper than that of type *3* through any given point in *Y,B*-space. That of type *2* is steeper by the difference in preference for leisure. That of type 1 is steeper by standard arguments[9] (see Stiglitz (1982)). It is a well-known property of the Pareto optimum that

$de^3 / dY^3 = 1$ (the zero marginal tax at the top). (29)

Adding the condition for Pareto optimality, we have five equations implicitly defining all other variables as functions of V^3 including $V^1(V^3)$, which can be interpreted as a Pareto frontier. Even if consumers of categories *1* and *2* have identical indifference maps and thus a utility function in *Y,B*-space with the same *ordinal* properties, the welfare weights that the government assigns to income changes for the two categories may be different and governed by some transformation of the ordinal utility function. We let the cardinalisation $V^2 = \nu(V^1)$, with $\nu'>0$, dictate the weight given to the high-skilled person with strong preference for leisure[10]. By letting $\nu'<1$, additional income for high-skilled with strong preference for leisure will be given less weight than additional income for low-skilled when the initial income point is the same.

We can now examine how the choice of tax policy will be affected by the composition of the low income group. We can then examine how the composition of the *m*-group affects the tax design. By increasing n^2 and lowering n^1 correspondingly we see how taxes are affected by the fact that those with a low income consists not only of low-skilled persons, but also of high-skilled with a strong preference for leisure. The welfare function to be maximized is the sum of welfare across the three groups:

$$\Omega = n^3 V^3 + n^2 V^2 + (m - n^2) V^1 \qquad (30)$$

[8] By substituting from (26) and (27) in (1) and equating the right hand sides of (27) and (28) we are left with two constraints and can maximise w.r.t. Y^1 and Y^3.

[9] The intuition is the following. Type 1, being less productive, has less leisure for a given value of *Y*. Hence he requires a larger compensation in terms of *B* for a marginal increase in labour supply. Being less productive, he also has to increase working hours more in order to generate a unit increase in gross income. This reinforces the conclusion that a higher increase in *B* is required to offset the disutility from a marginal increase in Y, and the slope of the indifference curve must be larger.

[10] A more general transformation might have been considered, but to avoid unnecessary complications we have chosen the simplest possible one for the purpose.

From the analysis above we can substitute $V^1(V^3)$ for V^1 and $V^2 = \nu(V^1(V^3))$ for V^2 and write the objective function as

$$\Omega = n^3 V^3 + n^2 \nu(V^1(V^3)) + (m - n^2) V^1(V^3) \tag{31}$$

Maximising w.r.t. V^3 we get the first order condition

$$\Omega' = N + n^2 \nu'(dV^1 / dV^3) + (m - n^2) dV^1 / V^3 = 0 \tag{32}$$

The second order condition is

$$\Omega'' < 0. \tag{33}$$

By comparative statics we can examine the effect of an increase in m^2 for a fixed m by differentiating the first order condition.

$$\Omega'' \frac{\partial V^3}{\partial n^2} + (\nu' - 1)(dV^1 / dV^3) = 0 \tag{34}$$

It follows that

$$\frac{\partial V^3}{\partial n^2} > 0 \text{ if } \nu' < 1. \tag{35}$$

The following proposition states the result.

Proposition 3: If low-skilled agents and high-skilled agents with strong preference for leisure, have identical indifference curves in *Y,B*-space, and if high-skilled agents with a strong preference for leisure are given a lower welfare weight than low-skilled individuals, a larger share of high-skilled agents in the low-income group motivates shifting more of the tax burden to the low-income group.

In other words the tax should be less redistributive if more of the low-income people are in fact not low-skilled, but have a low income as a consequence of working few hours.

The problem of distinguishing the low-skilled and the high-skilled with strong preference for leisure does not necessarily vanish even if people with different preferences have deviating indifference curves in *Y,B*-space since it may not be optimal to separate the two types at an income tax optimum. It may be optimal to have bunching of the low-skilled and the high-skilled with strong preference for leisure even though their indifference curves through the shared income point in *Y,B*-space have different slopes. If information on preference type were available, it might be

desirable to tax the high-skilled, leisure-prone type harder than the low-skilled, but in the absence of such information differentiated taxation may not be desirable.

The analysis above took as its point of departure the assumption of identical indifference curves of type *1* and type *2*. However, what we actually used in the analysis was the fact that $Y^2 = Y^1$ and $B^2 = B^1$, or in other words a bunching assumption. The latter is a less robust assumption as in the absence of identical indifference curves there may be bunching at the tax optimum for some values of V^3 but not for others. We then have to take into account that in the course of the comparative static analysis the bunching may be retained or abandoned. However we can state.

Proposition 4: If high-skilled agents with a strong preference for leisure are given a lower welfare weight than low-skilled individuals, a larger share of high-skilled agents in the low-income group motivates shifting more of the tax burden to the low-income group as long as there is bunching of low-skilled and high-skilled with strong preference for leisure.

4. What can commodity taxes achieve in the bunching case?

It is well known from the tax theory assuming homogeneity of preferences that an important role for commodity taxes may be to alleviate self-selection constraints of the kind considered above. (See for example Edwards et al. (1994)). This may be a role also in the case of heterogeneous preferences. A further role may be to provide a tool for taxing differently agents that the income tax is incapable of discriminating between. In fact this may happen also in the conventional theory, as there may be cases of bunching. However, we believe that the need is more urgent when preferences are heterogeneous as it is hard to distinguish between people who earn a low income because they are low-skilled, and people who choose a low income because they prefer to work short hours. This motivates us to consider the role of commodity taxes when there is bunching of the low-skilled type and the high-skilled with strong preference for leisure.

For expositional convenience we confine the price setting to the choice of a commodity tax for a single good. We denote by x^i the quantity of this good consumed by an individual of type *i*, and denote the consumer price by *p,* and the tax by *t.* The

pre-tax price is treated as fixed. To maximise welfare subject to the budget constraint and self-selection constraint barring type *3* from mimicking, we formulate the Lagrange function:

$$\Lambda = n^1 V^1\left(P, Y^1, B^1\right) + n^2 V^2\left(P, Y^2, B^2\right) + n^3 V^3\left(P, Y^3, B^3\right) - \beta\left(V^3(P, Y^1, B^1) - V^3(P, Y^3, B^3)\right)$$
$$+\mu\left(n^1\left(Y^1 - B^1\right) + n^2\left(Y^2 - B^2\right) + n^3\left(Y^3 - B^3\right) + t\sum_{i=1}^{3} x^i\right) \tag{36}$$

where P is the consumer price vector. Since we are considering the bunching case we impose the conditions $B^1 = B^2$ and $Y^1 = Y^2$ and rewrite the expression above as:

$$\Lambda = n^1 V^1\left(P, Y^1, B^1\right) + n^2 V^2\left(P, Y^1, B^1\right) + n^3 V^3\left(P, Y^3, B^3\right) - \beta\left(V^3(P, Y^1, B^1) - V^3(P, Y^3, B^3)\right)$$
$$+\mu\left(n^1\left(Y^1 - B^1\right) + n^2\left(Y^1 - B^1\right) + n^3\left(Y^3 - B^3\right) + t\sum_{i=1}^{3} x^i\right) \tag{37}$$

The policy decision is to choose B^1, Y^1, B^3, Y^3 and *t*. The first order conditions with respect to B^1, Y^1, B^3, Y^3 are:

$$\Lambda_{B^1} = n^1 V_B^1 + n^2 V_B^2 - \mu\left(n^1 + n^2\right) + \mu t x_B^1 n^1 + \mu t x_B^2 n^2 - \beta V_B^{31} = 0 \tag{38}$$

$$\Lambda_{Y^1} = n^1 V_Y^1 + n^2 V_Y^2 - \mu\left(n^1 + n^2\right) + \mu t x_Y^1 n^1 + \mu t x_Y^2 n^2 - \beta V_Y^{31} = 0 \tag{39}$$

$$\Lambda_{B^3} = n^3 V_B^3 - \mu n^3 + \mu t x_B^3 n^3 + \beta V_B^3 = 0 \tag{40}$$

$$\Lambda_{Y^3} = n^3 V_Y^3 + \mu n^3 + \mu t x_Y^3 n^3 + \beta V_Y^3 = 0 \tag{41}$$

Using Roy's theorem the first order condition with respect to *t* (*p*) can be written as:

$$\Lambda_p = n^1 V_B^1 x^1 - n^2 V_B^2 x^2 - n^3 V_B^3 x^3 + \mu\left(n^1 x^1 + n^2 x^2 + n^3 x^3\right) + \mu t x_p + \beta V_B^{31} x^{31} - \beta V_B^3 x^3 = 0 \tag{42}$$

We will manipulate these first order conditions to obtain a formula for the commodity tax rate. For this purpose the following expression will be useful

$$\Lambda_{B^2} = n^2 (V_B^2 - \mu(1 - t x_B^2)) \tag{43}$$

This expression is obtained by differentiating eq. (36) partially w.r.t. B^2. By the envelope theorem this is the welfare effect of a hypothetical increase in B^2, i.e. the welfare effect that would be obtainable if it were possible to undertake a partial increase in B^2. It reflects the desirability, but not the feasibility, of increasing the

after-tax income (decreasing the income tax) for type two. V_B^2 is the social benefit of increasing the after-tax income of a type two person, whereas $-\mu(1-tx_B^2)$ deducts the associated social cost, net of tax revenue, due to the commodity tax. In general we cannot tell the overall sign. However, if the situation is such that we would like to tax a type two person harder and redistribute the proceeds to a low skill person, we would believe Λ_{B^2} to be negative.

Letting s^i denote the substitution effect for the taxed good for individual *i,* defining $S = n^1 s^1 + n^2 s^2 + n^3 s^3$, and using the first order conditions (38)-(42) and the Slutsky decomposition, we can state:

Proposition 5: Assume the income tax optimum is characterized by bunching of the low-skilled type and the high-skilled with strong preference for leisure, then we would like to tax a commodity according to the rule

$$t = \frac{\Lambda_{B^2}(x^2 - x^1)}{\mu S} - \frac{\beta V_B^{31}(x^{31} - x^1)}{\mu S} \tag{44}$$

Proof: See Appendix B

We recognize the second part of this formula as the usual "mimicking term". The coefficient multiplying $(x^{31} - x^1)$ is positive (since $V_B^{31} > 0$, β and μ are positive Lagrange multipliers, and S is negative). Hence, the last term is positive if the "mimicker" consumes more of the good in question than does the low skill person. We should use the commodity tax to deter mimicking, thereby mitigating the self-selection constraint. Since the mimicker enjoys more leisure than the low skill person, the good should be taxed if the demand increases in leisure and subsidized if demand decreases in leisure. This is the same story as told in e.g. Edwards et al. (1994).

The first term, which we will denote the "redistribution term", is of another character. The difference in consumption between the (actual) high skill person with a strong preference for leisure and the consumption of the low skill person is of crucial importance. Λ_{B^2} shows how we value an increase in after-tax income to a type two person. Let us for the moment assume this term is negative. Then the term

$(\Lambda_{B^2} / \mu S)$ multiplying $(x^2 - x^1)$ is positive. If type two has a higher consumption than the low skill person, this term tends to increase the value of *t*. If the high-skilled type two is the larger consumer, one can impose a commodity tax and lower the income tax on the low-skilled individual to compensate him for the commodity tax burden. This tax relief will be inadequate to compensate the larger high-skilled consumer with strong preference for leisure. Thus an additional tax burden is imposed on exactly the person who was considered too leniently taxed in the first place.

To elaborate on the insights from formula (44), it is helpful to make use of the taxonomy in Section 2. Let us recall that the disposable income is the same for type 1, type 2 and the mimicker appearing in the formula. The hours of work is the same for the high skill person with strong preference for leisure and the mimicker, whereas the low skill person works more hours.

In case *i.* of the taxonomy market goods are separable from leisure in the utility function, and preferences for commodities are the same. People who have the same disposable income will therefore have the same demand. This implies that $x^1 = x^2 = x^{31}$ and that both the mimicking term and the redistribution term in (44) will vanish, and nothing can be gained by a commodity tax.

If the preferences are of the form in case *ii.,* that is $u^i(f(z,h),h)$, then the type two (actual) person and the (potential) mimicker will have the same consumption bundles, but the consumption for the low skill person will be different. If Λ_{B^2} is negative, the redistribution and mimicking terms will reinforce each other, and they both imply that there should be a commodity tax if demand increases in leisure and a subsidy if it decreases[11]. In this sense the two terms are similar. However, note the difference that no tax revenue is actually collected from the mimicker, whereas the type two person actually pays the tax. Even though we regard it as the normal case that Λ_{B^2} is negative, it might conceivably be positive, and the two terms in eq. (44) would counteract each other.

[11] A key assumption in Saez (2002) ensuring there is no need for commodity taxes is that conditional on income level the welfare weight assigned to additional income is uncorrelated to the consumption of individual goods. In our model it is clear that this condition is not satisfied if marginal income for the high-skilled is socially appreciated more highly than marginal income for the low-skilled when they choose the same income but different consumption bundles.

If preferences are as in case *iii.* in our taxonomy, i.e. preferences for type 3 and type 1 are of the form $U(f(z),h)$, but for type two persons of the form $U(\phi(z),h)$, then the low skill person and the mimicker will have the same consumption and the mimicker term in eq. (44) will vanish. However, the consumption of the high skill person with strong preference for leisure will differ from the consumption of the low skill person and the "redistribution" term will be nonzero. By exploiting the differences in consumption bundles one may impose commodity taxes on goods of which the high-skilled with strong preference for leisure have high consumption and possibly introduce subsidies favouring the low-skilled. Hence, even though we cannot use the income tax to redistribute from the high skill person with strong preference for leisure and the low skill person, we can achieve redistribution via the commodity tax.

If the utility function is not separable and preferences are heterogeneous, i.e. type 1 and type 3 have a utility function $U(z,h)$ and type 2 a utility function $\tilde{U}(z,h)$ the mimicking term would point in the direction of taxing the commodity if the demand according to $U(z,h)$ increases with leisure. However, even if $\tilde{U}(z,h)$ implies a demand function where x^2 is, say, increasing in leisure, it might still be the case that x^2 is less than x^1 and the redistribution term would motivate a subsidy.

5. Conclusion

We have highlighted the problems involved in tracing the implications of heterogeneous preferences for optimum taxes. In this context we have argued that it is helpful to classify various types of preference heterogeneity. Which case that prevails is shown to be crucial for what can be achieved by various tax instruments.

We have singled out two issues for further analysis. Firstly, we have characterised various income tax optima, showing how the treatment of people with different preferences will depend on the information constraints impeding income transfers. In particular we have discussed the case in which there is bunching of low-skilled individuals and high-skilled people with strong preference for leisure, and demonstrated how the optimum income tax will then depend on the composition of the population. We have also studied what assumptions are needed in order to rank individuals' utilities. Secondly, we have discussed the role of commodity taxes. We have identified the preference cases in which tax differentiation not attainable by

income taxes, can be achieved by means of commodity taxes. This may happen either because preferences for market commodities differ, or because high-skilled individuals have more leisure and leisure in turn induces consumption of certain commodities (leisure goods). Finally, we have discussed the interaction of preferences for market commodities and the preference for leisure, which implies that the trade-off between income and leisure in general will be affected by prices and hence can be manipulated by commodity taxes. Interestingly, this may provide a tool for separating individuals that would otherwise be bunched together at the income tax optimum, but only at a cost since commodity taxes are distortionary.

Appendix A

Distributional preferences.

We want to argue that for given individual preferences it will normally be possible to assume such social preferences that the same marginal social utility of income is assigned to both types of high-skilled individuals when they are taxed equally. Let $u^i(B^i, Y^i)$ be some ordinal utility function representing the individual preferences of type i (=2,3), and consider the case in which both types of high-skilled individuals pay a tax *T*, so that the utility of type *i* is $u^i(Y^i - T, Y^i)$. At an undistorted allocation u^i is maximised w.r.t. Y^i for a fixed *T*, and Y^i as well as the utility level can be perceived as functions of *T*. To get a proper representation of social preferences, we choose a suitable cardinalisation of each utility function by means of the monotonic transformation $f_i(u^i)$, where $f_i'(u^i) > 0$. V^i is then interpreted as the utility resulting from the selected cardinalisation, i.e. $V^i = f_i(u^i)$.

Assume that f_2 and f_3 can be chosen such that there is no case for transfers between the two types of high-skilled agents, i.e. for all values of *T* $\frac{d}{dT} f_2(u^2) = \frac{d}{dT} f_3(u^3)$ which, due to the envelope theorem, is equivalent to $f_2'(u^2)u_B^2 = f_3'(u^3)u_B^3$. The equation can be rewritten as:

$$f_2'(u^2) = \frac{u_B^3}{u_B^2} f_3'(u^3) \qquad \text{(A1)}$$

Normally we can find functions f_2 and f_3 (with $f_i'(u^i) > 0$) such that this condition will hold. Suppose in particular that a cardinalisation f_3 has been chosen to reflect the

distributional preferences with regard to people of type 3. With the functions u^2, u^3 and f_3 being known, the right hand side of (A1) becomes a function of *T*. Let us further invert the relationship between u^2 and *T*, and write *T* as a (strictly decreasing) function of u^2. We can then express the r.h.s. of (A1) as a function of u^2, called $\varphi(u^2)$, and write (A1) as $f_2'(u^2)=\varphi(u^2)$ which is a differential equation. For a wide class of functions this be solved to obtain f_2.

.

Appendix B.

Proof of proposition 5.

Making use of the Slutsky decomposition we can rewrite eq. (42) as :

$$\Lambda_p = -n^1V_B^1x^1 - n^2V_B^2x^2 - n^3V_B^3x^3 + \mu\left(n^1x^1 + n^2x^2 + n^3x^3\right) + \mu t\left(n^1s^1 + n^2s^2 + n^3s^3\right)$$
$$- \mu t\left(n^1x_B^1x^1 + n^2x_B^2x^2 + n^3x_B^3x^3\right) + \beta V_B^{31}x^{31} - \beta V_B^3x^3 = 0 \qquad \text{(B1)}$$

where s^i is the Slutsky derivative of x^i. Multiplying (38) by x^1, we get

$$n^1V_B^1x^1 - n^2V_B^2x^1 - \mu\left(n^1x^1 + n^2x^1\right) + \mu tx^1\left(n^1x_B^1 + x_B^2\right) - \beta V_B^{31}x^1 = 0 \qquad \text{(B2)}$$

Multiplying (40) by $-x^3$, yields

$$-n^3V_B^3x^3 + \mu n^3x^3 - \mu tx_B^3n^3x^3 - \beta V_B^3x^3 = 0 \qquad \text{(B3)}$$

We can make use of (B3) to eliminate a number of terms in (B1). Then adding each side of (B1) and (B2) we obtain

$$n^2V_B^2(x^1-x^2) - \mu n^2(x^1-x^2) + \beta V_B^{31}(x^{31}-x^1) + \mu tx_B^2(x^1-x^2)n^2 + \mu tS = 0 \qquad \text{(B4)}$$

where $S = n^1s^1 + n^2s^2 + n^3s^3$. Now returning to (36) and differentiating w.r.t. B^2, we find

$$\Lambda_{B^2} = n^2(V_B^2 - \mu(1-tx_B^2)) \qquad \text{(B5)}$$

We can then reformulate (B4) as

$$-\Lambda_{B^2}(x^2-x^1) + \beta V_B^{31}(x^{31}-x^1) + \mu tS = 0 \qquad \text{(B6)}$$

which can be written as eq (44).

References

Atkinson, A.B. and J.E. Stiglitz (1976), The Design of Tax Structure: Direct versus Indirect Taxation, *Journal of Public Economics* 6, 55-75.

Blomquist, S. and L. Micheletto (2003). Age Related Optimal Income Taxation, Working Paper 2003:7, Department of Economics, Uppsala University.

Boadway R.; Marchand M.; Pestieau P., del Mar Racionero M. (2002): Optimal Redistribution with Heterogeneous Preferences for Leisure, *Journal of Public Economic Theory* 4 (no. 4), 475-498

Cuff, K. (2000). Optimality of Workfare with Heterogeneous Preferences. *Canadian Journal of Economics 33,* 149-174.

Edwards, J., Keen M., and Tuomala M. (1994). Income Tax, Commodity Taxes and Public Good Provision: A Brief Guide. *FinanzArchiv 51*, 472-487.

Jordahl, H. and L. Micheletto (2002). "Optimal Utilitarian Taxation and Horizontal Equity." Working paper 2002:19, Department of Economics, Uppsala University, forthcoming in *Journal of Public Economic Theory*

Mirrlees, J.A. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies* 38, 175-208.

Nozick, R. (1974). *Anarchy, State, and Utopia*. Oxford: Blackwell.

Saez, E. (2002). The Desirability of Commodity Taxation under Non-Linear Income Taxation and Heterogeneous Tastes, *Journal of Public Economics 83*, 217-230.

Sandmo, A. (1993) : Optimal Redistribution when Tastes Differ, *Finanz Archiv 50*(2), 149-163.

Stern, N.H. (1982). Optimum Income Taxation with Errors in Administration. *Journal of Public Economics 17*, 181-211.

Stiglitz, J.E. (1982). Self-Selection and Pareto Efficient Taxation. *Journal of Public Economics 17*, 213-240.

Tarkianen, R. and M. Tuomala (1999). Optimal Nonlinear Income Taxation with a Two-Dimensional Population; A Computational Approach. *Computational Economics 13*, 1-16.

CESifo Working Paper Series

(for full list see www.cesifo.de)

1178 Stefan Lachenmaier and Ludger Woessmann, Does Innovation Cause Exports? Evidence from Exogenous Innovation Impulses and Obstacles, April 2004

1179 Thiess Buettner and Johannes Rincke, Labor Market Effects of Economic Integration – The Impact of Re-Unification in German Border Regions, April 2004

1180 Marko Koethenbuerger, Leviathans, Federal Transfers, and the Cartelization Hypothesis, April 2004

1181 Michael Hoel, Tor Iversen, Tore Nilssen, and Jon Vislie, Genetic Testing and Repulsion from Chance, April 2004

1182 Paul De Grauwe and Gunther Schnabl, Exchange Rate Regimes and Macroeconomic Stability in Central and Eastern Europe, April 2004

1183 Arjan M. Lejour and Ruud A. de Mooij, Turkish Delight – Does Turkey's accession to the EU bring economic benefits?, May 2004

1184 Anzelika Zaiceva, Implications of EU Accession for International Migration: An Assessment of Potential Migration Pressure, May 2004

1185 Udo Kreickemeier, Fair Wages and Human Capital Accumulation in a Global Economy, May 2004

1186 Jean-Pierre Ponssard, Rent Dissipation in Repeated Entry Games: Some New Results, May 2004

1187 Pablo Arocena, Privatisation Policy in Spain: Stuck Between Liberalisation and the Protection of Nationals' Interests, May 2004

1188 Günter Knieps, Privatisation of Network Industries in Germany: A Disaggregated Approach, May 2004

1189 Robert J. Gary-Bobo and Alain Trannoy, Efficient Tuition Fees, Examinations, and Subsidies, May 2004

1190 Saku Aura and Gregory D. Hess, What's in a Name?, May 2004

1191 Sjur Didrik Flåm and Yuri Ermoliev, Investment Uncertainty, and Production Games, May 2004

1192 Yin-Wong Cheung and Jude Yuen, The Suitability of a Greater China Currency Union, May 2004

1193 Inés Macho-Stadler and David Pérez-Castrillo, Optimal Enforcement Policy and Firms' Emissions and Compliance with Environmental Taxes, May 2004

1194 Paul De Grauwe and Marianna Grimaldi, Bubbles and Crashes in a Behavioural Finance Model, May 2004

1195 Michel Berne and Gérard Pogorel, Privatization Experiences in France, May 2004

1196 Andrea Galeotti and José Luis Moraga-González, A Model of Strategic Targeted Advertising, May 2004

1197 Hans Gersbach and Hans Haller, When Inefficiency Begets Efficiency, May 2004

1198 Saku Aura, Estate and Capital Gains Taxation: Efficiency and Political Economy Consideration, May 2004

1199 Sandra Waller and Jakob de Haan, Credibility and Transparency of Central Banks: New Results Based on Ifo's *World Economicy Survey*, May 2004

1200 Henk C. Kranendonk, Jan Bonenkamp, and Johan P. Verbruggen, A Leading Indicator for the Dutch Economy – Methodological and Empirical Revision of the CPB System, May 2004

1201 Michael Ehrmann, Firm Size and Monetary Policy Transmission – Evidence from German Business Survey Data, May 2004

1202 Thomas A. Knetsch, Evaluating the German Inventory Cycle – Using Data from the Ifo Business Survey, May 2004

1203 Stefan Mittnik and Peter Zadrozny, Forecasting Quarterly German GDP at Monthly Intervals Using Monthly IFO Business Conditions Data, May 2004

1204 Elmer Sterken, The Role of the IFO Business Climate Indicator and Asset Prices in German Monetary Policy, May 2004

1205 Jan Jacobs and Jan-Egbert Sturm, Do Ifo Indicators Help Explain Revisions in German Industrial Production?, May 2004

1206 Ulrich Woitek, Real Wages and Business Cycle Asymmetries, May 2004

1207 Burkhard Heer and Alfred Maußner, Computation of Business Cycle Models: A Comparison of Numerical Methods, June 2004

1208 Costas Hadjiyiannis, Panos Hatzipanayotou, and Michael S. Michael, Pollution and Capital Tax Competition within a Regional Block, June 2004

1209 Stephan Klasen and Thorsten Nestmann, Population, Population Density, and Technological Change, June 2004

1210 Wolfgang Ochel, Welfare Time Limits in the United States – Experiences with a New Welfare-to-Work Approach, June 2004

1211 Luis H. R. Alvarez and Erkki Koskela, Taxation and Rotation Age under Stochastic Forest Stand Value, June 2004

1212 Bernard M. S. van Praag, The Connexion Between Old and New Approaches to Financial Satisfaction, June 2004

1213 Hendrik Hakenes and Martin Peitz, Selling Reputation When Going out of Business, June 2004

1214 Heikki Oksanen, Public Pensions in the National Accounts and Public Finance Targets, June 2004

1215 Ernst Fehr, Alexander Klein, and Klaus M. Schmidt, Contracts, Fairness, and Incentives, June 2004

1216 Amihai Glazer, Vesa Kanniainen, and Panu Poutvaara, Initial Luck, Status-Seeking and Snowballs Lead to Corporate Success and Failure, June 2004

1217 Bum J. Kim and Harris Schlesinger, Adverse Selection in an Insurance Market with Government-Guaranteed Subsistence Levels, June 2004

1218 Armin Falk, Charitable Giving as a Gift Exchange – Evidence from a Field Experiment, June 2004

1219 Rainer Niemann, Asymmetric Taxation and Cross-Border Investment Decisions, June 2004

1220 Christian Holzner, Volker Meier, and Martin Werding, Time Limits on Welfare Use under Involuntary Unemployment, June 2004

1221 Michiel Evers, Ruud A. de Mooij, and Herman R. J. Vollebergh, Tax Competition under Minimum Rates: The Case of European Diesel Excises, June 2004

1222 S. Brock Blomberg and Gregory D. Hess, How Much Does Violence Tax Trade?, June 2004

1223 Josse Delfgaauw and Robert Dur, Incentives and Workers' Motivation in the Public Sector, June 2004

1224 Paul De Grauwe and Cláudia Costa Storti, The Effects of Monetary Policy: A Meta-Analysis, June 2004

1225 Volker Grossmann, How to Promote R&D-based Growth? Public Education Expenditure on Scientists and Engineers versus R&D Subsidies, June 2004

1226 Bart Cockx and Jean Ries, The Exhaustion of Unemployment Benefits in Belgium. Does it Enhance the Probability of Employment?, June 2004

1227 Bertil Holmlund, Sickness Absence and Search Unemployment, June 2004

1228 Klaas J. Beniers and Robert Dur, Politicians' Motivation, Political Culture, and Electoral Competition, June 2004

1229 M. Hashem Pesaran, General Diagnostic Tests for Cross Section Dependence in Panels, July 2004

1230 Wladimir Raymond, Pierre Mohnen, Franz Palm, and Sybrand Schim van der Loeff, An Empirically-Based Taxonomy of Dutch Manufacturing: Innovation Policy Implications, July 2004

1231 Stefan Homburg, A New Approach to Optimal Commodity Taxation, July 2004

1232 Lorenzo Cappellari and Stephen P. Jenkins, Modelling Low Pay Transition Probabilities, Accounting for Panel Attrition, Non-Response, and Initial Conditions, July 2004

1233 Cheng Hsiao and M. Hashem Pesaran, Random Coefficient Panel Data Models, July 2004

1234 Frederick van der Ploeg, The Welfare State, Redistribution and the Economy, Reciprocal Altruism, Consumer Rivalry and Second Best, July 2004

1235 Thomas Fuchs and Ludger Woessmann, What Accounts for International Differences in Student Performance? A Re-Examination Using PISA Data, July 2004

1236 Pascalis Raimondos-Møller and Alan D. Woodland, Measuring Tax Efficiency: A Tax Optimality Index, July 2004

1237 M. Hashem Pesaran, Davide Pettenuzzo, and Allan Timmermann, Forecasting Time Series Subject to Multiple Structural Breaks, July 2004

1238 Panu Poutvaara and Andreas Wagener, The Invisible Hand Plays Dice: Eventualities in Religious Markets, July 2004

1239 Eckhard Janeba, Moral Federalism, July 2004

1240 Robert S. Chirinko, Steven M. Fazzari, and Andrew P. Meyer, That Elusive Elasticity: A Long-Panel Approach to Estimating the Capital-Labor Substitution Elasticity, July 2004

1241 Hans Jarle Kind, Karen Helene Midelfart, Guttorm Schjelderup, Corporate Tax Systems, Multinational Enterprises, and Economic Integration, July 2004

1242 Vankatesh Bala and Ngo Van Long, International Trade and Cultural Diversity: A Model of Preference Selection, July 2004

1243 Wolfgang Eggert and Alfons J. Weichenrieder, On the Economics of Bottle Deposits, July 2004

1244 Sören Blomquist and Vidar Christiansen, Taxation and Heterogeneous Preferences, July 2004