SANNA TENHUNEN

Essays on the Theory of Optimal Taxation

■

ACADEMIC DISSERTATION
To be presented, with the permission of
the Faculty of Economics and Administration
of the University of Tampere, for public discussion in the
Auditorium Pinni B 1097 of the University, Kanslerinrinne 1,
Tampere, on December 7th, 2007, at 12 o'clock.

UNIVERSITY OF TAMPERE

ACADEMIC DISSERTATION
University of Tampere
Department of Economics and Accounting
Finland

Distribution
Bookshop TAJU
P.O. Box 617
33014 University of Tampere
Finland

Tel. +358 3 3551 6055
Fax +358 3 3551 7685
taju@uta.fi
www.uta.fi/taju
http://granum.uta.fi

Cover design by
Juha Siro

Acta Universitatis Tamperensis 1276
ISBN 978-951-44-7139-1 (print)
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 672
ISBN 978-951-44-7140-7 (pdf)
ISSN 1456-954X
http://acta.uta.fi

Tampereen Yliopistopaino Oy – Juvenes Print
Tampere 2007

Acknowledgements

During the last five years of writing this doctoral thesis I have received help, assistance and support from many people. I would like to thank Professor Matti Tuomala, co-author and supervisor, and Dr. Jukka Pirttilä, co-author and unofficial supervisor. They have been advising, guiding and encouraging me at every point of this process, all the way from the times I was an undergraduate. They never seemed to lose their faith on me and my abilities. They have been reading hundreds of pages of my writing, equations and tables, commented, made suggestions and discussed ideas to improve my work. I deeply appreciate all the work you have done.

I have benefited enormously from the insightful comments and suggestions received from Thomas Gaube, Vidar Christiansen and the other participants at FDPE workshops, international conferences and other seminars. I would also like to thank the pre-examiners, Professors Thomas Aronsson and Panu Poutvaara for their careful work and comments that helped me to improve this thesis considerably. I am also grateful to Virginia Mattila for helping me with the language.

The study was carried out with financial support from the Yrjö Jahnsson Foundation, the Department of Economics and Accounting, University of Tampere and the graduate school fellowship provided by the Finnish Doctoral Programme in Economics. I am grateful for the opportunity to concentrate full-time on research and postgraduate studies.

I would also like to thank all my colleagues at the Department of Economics and Accounting for the inspiring atmosphere, encouragement and help I have received during these years. Especially I would like to thank Kaisa who, in addition for being my friend, has helped me a lot in both technical and practical challenges in studies and research. I would also like to thank Markus for the advice and support he has given me during these years. His enthusiasm and optimism in planning the future gave me faith: things tend to get organised! My warmest thanks for discussions on economics and non-economics go also to Harri, Saikat, Jani-Petri and Elina, likewise Päivi and Petri.

Finally, I would like to thank all my friends and family, who have had the unpleasant task of cheering me up when I was desperate and pulling my feet back onto the ground when the scientific world started to play too big a role. I am grateful to my parents, Sirkka ja Jussi, and my brother Juho, who have always been there for me when I have needed someone to talk to. And most of all, I would like to thank my dear husband, Petteri, who has shown enormous patience, love, belief and understanding during these years. I love you, too.

Helsinki, October 2007

Sanna Tenhunen

Copyright notice

Abstract

Taxation is a powerful device for affecting the behaviour of consumers. This is both an advantage, in the case of using taxes to discourage, or subsidies to encourage the consumption of some good, but also a disadvantage, as a result of the negative incentive effects that redistributive taxation creates. The justification for taxation might be purely efficiency, when taxation is unambiguously welfare improving. But when the justification for tax policy is based on considerations of equality or paternalistic objectives, the optimality of such taxation becomes more a question of values.

Tax policy, in its widest sense, includes almost all public sector activities: both public sector finance in terms of tax revenues and expenditures like social security and public provision are a part of tax policy. A theoretical consideration of so large an entity is complex, and thus it is reasonable to break down the problem into smaller parts. Although the setting may not correspond to real life, it allows a more exact handling of the details building a bigger picture than would otherwise be possible. The purpose of this thesis is to take gradual steps towards a more general setting in the theoretical framework of optimal taxation. The basic framework in the background considers the trade-off between efficiency and equality in a case where individuals differ in their wage earning abilities and that there is asymmetric information on this characteristic restricting the choice on the optimal tax policy. In this thesis the framework is generalised to a world with multidimensional heterogeneity among agents in several settings.

The two first essays respectively consider the use of commodity taxes together with income taxation in a model with environmental objectives in a world with non-fixed wages and heterogeneous preferences. The possible trade-off between environmental and redistributive objectives, suggested by the earlier empirical findings of regressive environmental taxes, is considered in both essays: it is

found that that some of the harm from the externality may be compensated by gains in redistribution. It is also found that the efficiency and equity based parts in the commodity tax rates cannot necessarily be separated from each other.

The third essay considers paternalistic objectives and influencing the consumption of a meritorious good with the help of public provision in a case where some of the agents are not perfectly rational. It is found that even if both the aim to reduce the harm from irrationality and redistributive income taxation separately would speak for a positive level of public provision, when these two aims are put together, it is no longer clear that public provision improves welfare.

The last essay concerns life-time redistribution in a case where agents have different discount rates. It concentrates on the tax treatment of savings as a device to affect life-time distribution. An alternative interpretation is a pension programme. The role of taxation of savings is considered in both a theoretical and a numerical framework. It is found that in a world with multidimensional heterogeneity there is a case for the non-linear taxation of savings.

Tiivistelmä

Verotus on tehokas työkalu kulutuskäyttäytymisen ohjaamiseen, mutta kohdistuessaan eri tavalla eri henkilöryhmiin sillä on myös kielteisiä kannustinvaikutuksia. Esimerkiksi tilanteessa, jossa kulutuksella on haitallisia ulkoisvaikutuksia, verotuksen avulla voidaan parantaa kaikkien asemaa. Veroratkaisun taustalla voi olla myös oikeudenmukaisuuskäsitykseen tai paternalismiin perustuvia syitä, jolloin yksilön näkökulmasta verotus tuottaa sekä voittajia että häviäjiä. Tällaisissa tilanteissa verotuksen optimaalisuus on arvovalinnoista riippuva kysymys.

Verojärjestelmä laajasti ymmärrettynä kattaa lähes koko julkisen sektorin toiminnan: julkisen sektorin rahoituksen pohjan muodostavien verotulojen lisäksi verojärjestelmään voidaan laskea kuuluvaksi myös julkisia menoja, kuten sosiaaliturva, tulonsiirrot ja julkinen tuotanto. Näin laajan ja monimutkaisia vuorovaikutuksia sisältävän kokonaisuuden teoreettinen tarkastelu on järkevää suorittaa pienempinä palasina, jolloin myös yksityiskohtiin voidaan kiinnittää tarkemmin huomiota. Tässä väitöskirjassa pyritään ottamaan askeleita kohti yleisempää ja todenmukaisempaa mallikehikkoa optimaalisen verotuksen teoriassa. Lähtökohtana on malli, jossa tarkastellaan verotuksen tehokkuus- ja oikeudenmukaisuusnäkökulmia ja niiden yhdistämistä. Taustaoletuksena on, että yksilöillä on erilaiset tulonansaintamahdollisuudet, joita verotuksen suunnittelija ei pysty havaitsemaan. Tämä informaation epäsymmetria rajoittaa optimaalisen verotuksen suunnittelua. Tutkimus koostuu neljästä esseestä, joissa mallia laajennetaan realistisempaan suuntaan luopumalla oletuksesta, että tulonansaintamahdollisuudet ovat ainoa ominaisuus, minkä suhteen yksilöt eroavat.

Kaksi ensimmäistä esseetä käsittelevät hyödykeverotuksen käyttöä tilanteessa, missä kulutuksella on haitallisia ympäristövaikutuksia. Esseissä tarkastellaan teoreettisen kehikon avulla ympäristö- ja tulonjakotavoitteiden välistä yhteyttä

aikaisemman ympäristöverotuksen regressiivisyydestä viitteitä antaneen kirjallisuuden innoittamana. Tulosten mukaan näillä tavoitteilla on vastakkaisiin suuntiin osoittavia piirteitä eikä hyödykeverotuksen tehokkuuteen ja oikeudenmukaisuuteen pohjaavia osia voida aina erotella.

Kolmannessa esseessä käsitellään julkisen sektorin tuotannon optimaalisuutta tilanteessa, jossa kaikki kuluttajat eivät ole täysin rationaalisia ja verotuksen suunnittelun taustalla on paternalistisia tavoitteita. Tutkimuksessa osoitetaan, että vaikka sekä paternalistiset tavoitteet että uudelleenjakava tuloverotus yksistään puoltavat julkista tuotantoa, yhdistettäessä nämä kaksi tavoitetta julkisen tuotannon hyvinvointia lisäävä vaikutus hämärtyy.

Neljännessä esseessä tarkastellaan teoreettisen lähtökohdan lisäksi myös numeerisesti eliniän sisäistä tulonjakoa tilanteessa, jossa yksilöt diskonttaavat tulevaisuutta eri kertoimilla. Essee keskittyy tarkastelemaan säästämisen verokohtelua työkaluna eliniän kulutusmahdollisuuksien tasaamiseksi. Vaihtoehtoisesti malli voidaan tulkita myös eläkejärjestelmäksi. Tulosten mukaan säästämisen epälineaarinen verotus on optimaalista tilanteessa, jossa yksilöt eroavat sekä diskonttokertoimen että tulonansaintakyvyn suhteen.

Table of contents

Chapter 1
Essays on the Theory of Optimal Taxation: An Overview

Introduction

Taxation is always a topical research topic, not least because of its policy relevance. The design of an optimal tax policy, in a wide sense including all public services funded with tax revenue, affects all members of a society. Because of informational restrictions on individuals' abilities, needs or preferences, the trade-off between efficiency and equity objectives plays an important role in the design of optimal tax policy.

This thesis consists of four independent essays, with optimal taxation as the common factor. In this chapter there is first an introduction to what is meant by optimal taxation. There is also a short review of the history of taxation and tax policy, discussion of the current state of the literature on optimal taxation and finally, a summary of the essays.

On the concept of optimal taxation

A phrase 'optimal taxation' imposes high demands on the research. Has somebody solved what taxes should optimally be? Unfortunately, the answer is 'no'. Neither have the earlier contributions on theory done so nor will the essays in this thesis be able to meet such a challenge. Optimality of taxes is always, at least to some extent, a question of values. For example, income redistribution, a very commonly used justification for taxation, is a widely debated issue with not just one correct solution. Thus by the means of positive economics optimal tax policy cannot be determined explicitly.

Another easily emerging misunderstanding to be rectified here is the scope of the term ‘taxation’. A narrow interpretation for taxation includes only those payments levied by the state or municipalities, such as income tax or value added tax. But other transfers that have to be paid, such as social security contributions can also be included in taxation. Extending the interpretation further, the way in which taxes are used should also be considered when talking about taxation in general. Thus subsidies and publicly provided goods are also part of tax policy. Taking a step towards a wider concept of taxation, the whole wedge between gross wage and purchasing power can be called an implicit tax.

Instead of solving the optimal levels of taxes the essays in this thesis discuss some rules for tax policy that can be derived in a theoretical microeconomic framework. These rules concern almost without exception only marginal taxes, i.e. how an additional unit of the income, output or other monetary outcome subjected to taxation should be treated. In microeconomics marginal concepts have been used as tool in determining the choice of consumers and firms ever since neoclassical economics was introduced at the end of the 19th century. In spite of lack of measurability and observability, the marginal concepts have asserted their contribution to decision-making.

Some history of the theory of optimal taxation

The theory of optimal taxation depends heavily on the philosophical foundations of political economy. The origins of these ideas can be traced back to Plato’s notion of a perfect state, as taken to a starting point in Schumpeter (1954) considering the history of economics in detail. However, in this brief overview of the history of the economics of taxation we concentrate on the more contemporary literature, starting from the latter part of the 18th century with Adam Smith's *Wealth of Nations*[1], where modern economic science can be thought to have begun. Although the emphasis here is on the discussion of the optimality of taxes, it might be interesting first to highlight some historical points on the practise of taxation.[2]

The commencement of taxation

The exact time when the first taxes were imposed is not known, but they have been collected at least ever since the time of Egyptian Pharaohs and Mesopotamian communities. At that time taxes were mostly collected in kind, based on easily observable characteristics such as ownership of production factors like land or slaves.[3] In the feudal period there was a predecessor for income tax, *scutage*, a payment allowing knight to avoid military service. As wars, especially lost wars, required revenues, there were strong incentives to raise taxes. As a result a legitimacy of taxation was needed. One of the first was *Magna Carta* issued in 1215, stating that increasing tax or imposing a new tax requires the consent of Parliament.

As economy transformed from agrarian to monetary-based economy and the property income of the state ceased with a decline of the feudal system, there was

[1] Adam Smith's "*An Inquiry into the Nature and Causes of the Wealth of Nations*", first published in 1776, can be read in modernised English, for example at http://www.adamsmith.org/smith/won-index.htm.

[2] The discussion of the early history of taxation and economic thought is mainly based on Schumpeter (1959), Musgrave (1985), Salanié (2003) and Screpanti and Zamagni (2005).

[3] Hudson (2000) provides an introduction to taxation in ancient times (with an emphasis on the land-value taxation).

a need for modern type taxation. A first modern income tax was imposed in the late 18th century in Britain to finance the Napoleonic wars. Later, income tax was abolished and reintroduced a few times, until it was permanently established in Britain in 1842, other countries following the example.

In the late 19th century the atmosphere was warming for social aspects. A good example of that was Prussia, where a compulsory health insurance and a pension system were introduced in the 1880's. With a rising interest in the social aspects the optimality of taxes started to include a question of tax incidence; who should pay taxes and how much?

Discussion on optimal taxes

The main problem of optimal taxation is the trade-off between efficiency and equity. Adam Smith in *Wealth of Nations* discussed at length what taxation should be like. In addition to writing an analysis of the various taxes on rent, profit and wages in use at that time, he presented four principles regarding taxes in general:

- Equality aspect: taxes should be paid in proportion to tax payers' ability
- The time, manner and quantity of tax payment should be clearly defined, not arbitrary
- Taxes should be collected at a time and in a manner that is most likely to be convenient to payer
- Administrative costs should be kept as low as possible.

In general these principles still hold in modern economies, although paying taxes in kind (referring to the manner of tax payment) is no longer a relevant question.

Moral philosophers at the end of the 19th century discussed the optimality and the incidence of taxation. The idea of the greatest sum of happiness as a criterion for taxation was introduced along with the utilitarian concept. John Stuart Mill determined that the sacrifices required by taxation should put as equal pressure as possible upon all taxpayers. This can be interpreted as a rule for ‘optimal’ taxation in the sense that it fulfils the demands of vertical equity (equal treatment of people with unequal observable characteristics, in this case taxpayer’s ability to pay). This rule suggests that the tax burden should be distributed so that the rich pay higher sums in taxes than the poor.

Mill’s follower, F. Y. Edgeworth (1897) also considered the distribution of tax burden among taxpayers. He suggested that the marginal disutility should be the same for each taxpayer. This principle of equal marginal sacrifice together with diminishing marginal utility of income implies progressive taxation. Edgeworth noted that the sum of utilities would increase as the income distribution becomes more and more equal, until full egalitarianism is achieved. However, the result was known to be problematic, as was already brought up by Sidgwick (1883). He stated that greater equality leads to an increase in population, threatens liberty, variety and diversity of opinions and tastes and decreases the total amount of income as a result of greater preference for leisure. The last part of his warning can be interpreted to refer to a contemporary problem, incentives.

In the 20th century there seemed to be quite a large consensus that a fair tax burden ought to be progressive. There was some effort to find an optimal tax schedule and an optimal rate of progression, put forward by e.g. Cassel (1901) and Edgeworth (1919). Pigou (1928) examined tax rules applicable to several equity rules and found, following Edgeworth (1897), the principle of an equal marginal sacrifice and progressive taxation to be optimal.

In the 1930s the assumption underlying the results of the optimality of progressive taxation, the feasibility of interpersonal utility comparisons, was questioned (Robbins, 1932, 1938). It was argued that as long as the optimality of the tax schedule cannot be determined as Pareto efficient, the question of tax burden remains to be a matter of ethics or politics, but cannot be solved within economics. The importance of the question of the distribution of the fruits and burdens in the economy brought the subject back to the literature along with a discussion on social welfare functions (Bergson, 1938).

Even if the optimality of progressive taxation seemed to be widely accepted, the efficiency losses in production resulting from the negative incentive effects were also recognised. Pigou’s (1947) view of the trade-off between the redistributive aspect of taxation and incentive problem in production was that the redistribution from the rich to the poor is welfare improving (as a result of decreasing marginal utility of income) as long as it does not decrease the aggregate level of output. Corlett and Hague (1953) considered also the excess burden of taxation and compared direct and indirect taxation. They stated that goods more complementary to leisure should be taxed more heavily in order not to discourage labour supply.

The modern theory of optimal commodity and income taxation

The modern theory of optimal taxation that is still in the background of current models was developed in the 20th century. The theory of different types of taxes diverged from one common theory as the models considered became more detailed. This Section introduces some of the most influential results in commodity and income taxation that are essential to the frameworks used in the essays of this thesis.[4]

[4] As a result of the objective to concentrate on the issues that are central to the essays of this thesis, an important branch of taxation, corporate taxes and capital income taxes, has been deliberately excluded from this section with the exception of a short note on the taxation of savings.

Commodity taxation

The theory of optimal commodity taxation took a step forward when Ramsey (1927) introduced a rule for optimal commodity taxes.[5] He stated that instead of a uniform tax rate on all goods, taxes should be such that introducing them reduces the production of each taxed commodity in the same proportion. In the special case where the demand for goods is independent, this rule condenses to the 'inverse-elasticity rule': the higher the elasticity, the lower the tax. Ramsey's rule for commodity taxation was generalized to a many-consumer economy by Diamond and Mirrlees (1971), when they abandoned the unrealistic assumption of the availability of lump-sum taxes and transfers. The theory of optimal commodity taxation was developed further by discussions on the tax rules e.g. by Stiglitz and Dasgupta (1971), considering the restrictions of the available taxes, Diamond (1975), combining commodity taxes with redistributional objectives, and Atkinson and Stiglitz (1972) and Deaton (1977), calculating optimal commodity taxes under different assumptions.

In addition to the efficiency perspective on commodity taxation, its role as a correcting instrument was brought up. Marshall had already noticed that there is a justification for government to intervene in the market outcome to correct its distortions. Pigou (1920) introduced a formal model suggesting that when there are externalities, i.e. the social costs of consumption differ from the private costs, a government intervention in the form of taxes or subsidies is beneficial. The analysis of external costs has been widely applied to environmental economics, where corrective taxes can be used to struggle with negative externalities like the consumption of polluting goods. Positive externalities in turn can be applied to the analysis of merit goods in the sense introduced in Musgrave (1959). Further

[5] A formally identical result was obtained independently by Boiteux (1956) as a solution to the optimal pricing problem for a monopoly.

applications to the world with negative externalities are presented by Sandmo (1975) and Dixit (1985) considering the optimal tax treatment of goods with negative externalities. The optimal taxation of merit goods has been considered e.g. in Besley (1988).

Yet another important application of commodity taxation is the taxation of savings. In a dynamic model the consumption of the future period can be interpreted just like an additional commodity. Thus, the results of optimal commodity taxes can easily be applied to the problems of taxation of savings or pension programmes (as forced saving). However, there are some special features in dynamic models that are likely to distinguish future consumption from other commodities. One such is individuals' tendency to save too little. The optimal tax treatment of savings or capital income in general is considered in Ordover and Phelps (1979) presenting the zero capital tax result, Atkinson and Sandmo (1980) finding that the optimality of taxing capital depends on the policy instruments of the capital, and Feldstein (1985) considering the social security aspect of taxation of savings.

Income taxation

A hint of the modern way to consider the compromise between equity and the incentive effects of progressive taxation was given in Vickrey (1945), who was the first to formulate the question as a calculus of variation problem. After him, there was mostly only discussion on the effects of imperfect information on optimal taxes, e.g. by Graaff (1957). Some analysis of the optimal tax was presented by Zeckhauser (1970), Wesson (1972) and Fair (1971), but the formal modelling of the problem was introduced by the seminal work of James Mirrlees (1971) proposing a technique to consider the trade-off between efficiency and equity created by distortive taxation.

The essential assumption in Mirrlees' framework is that individuals have different abilities reflecting on their wages. Having higher wage earning ability puts some individuals in a potentially better situation, thus gives reasoning for redistributive taxation. Individuals can still choose if they want to take the advantage of this opportunity or not. The government is unable to detect either ability or the labour supply but can only monitor individual's income. This lack of perfect information restricts the design of a redistributive tax schedule: if the tax for the high-ability individuals is sufficiently high, all of them will supply only that amount of labour that gives them the same status as the low-ability individuals facing lower taxation. The main result of Mirrlees' work is that the optimal taxes are non-linear. This non-linearity results from the asymmetry of information, not from the different productivities.

The work of Mirrlees presents a technically advanced approach to the theory of optimal taxation. Once the method for an information constrained optimal tax analysis was developed, was been applied to all kinds of optimal tax problems. The most advanced analysis was presented by Mirrlees himself (Mirrlees, 1976), where he applied the method to find tax rules for linear and non-linear taxation, for an integrated tax system with a mix of income and commodity taxes and for optimal provision of public goods. The fundamental parts of the theory of optimal taxation, with an emphasis on mathematical problems were considered in Mirrlees (1986).

The model considering optimal non-linear income taxation includes both technical and practical difficulties in choosing an entire tax function for a continuum of people. Thus Mirrlees' successors have widely adopted a simplified version of the model with a discrete distribution of types to gain a more profound comprehension of the mechanism behind the results. The discrete approach was first captured in four different papers at the same time; Stiglitz (1982) and Stern (1982) considered the non-linear tax question, and with a slightly different emphasis Guesnerie and Seade (1982) examined tax distortions in general with a model of nonlinear pricing, and Nichols and Zeckhauser (1982) studying the restrictions the incentive effect imposes on redistribution.

Optimal income taxation has been considered in Mirrlees's footsteps with a variety of different assumptions. Research on optimal income tax has been conducted e.g. by Sheshinski (1972), Stern (1976) and Dixit and Sandmo (1977) with linear income tax and Stern (1976) for numerical calculation, Atkinson and Stiglitz (1976) in a case with non-linear income tax and linear commodity tax and Tuomala (1990) with several applications to the model concerning the redistributive effects of taxation.

Current trends in research field of optimal taxation

Despite its long history, the structure of optimal taxation remains a famous research topic in public economics. Even though much progress has been made, the models used continue to be simplifying characterisations of reality. In most optimal tax models competitive behaviour of firms is assumed.[6] There is also a lively discussion on the role of the governments' preferences, how their objectives are formed, as well as on the decentralisation of the public sector's tasks. The natural trend in research is to develop models towards more plausible assumptions of agents' characteristics and their behaviour. This has also led to more complicated models and a need for numerical characterisation of the problem. The whole branch of the literature on optimal taxation has become more complex and

[6] There is some literature on the issue with other market structures. For example, Fuest and Huber (1997) and Aronsson and Sjögren (2003, 2004) have considered optimal taxation with imperfect competition in the labour markets, and Auerbach and Hines (2001) provide a survey of optimal taxation and imperfect competition in general.

has diverged in several directions. Thus, here only a few research trends are presented that are most closely related to the topics of this thesis.[7]

A criterion for optimal taxation: from utilitarianism to paternalism

The choice of the social welfare function that the planner maximises when designing the tax policy naturally affects the results on the optimal taxation. This choice is largely a normative question. The justification for different social welfare functions could be retrieved just as well from ethics or philosophy as from economic efficiency. [8]

The least arguable criterion for taxation is Pareto efficiency, implying that such changes where at least one person gains while nobody loses are not available. In a world with asymmetric information on the taxpayers' characteristics non-distorting lump-sum taxes are not an option. Thus taxes always lead to efficiency losses implying that somebody has to be worse off and the Pareto criterion does not work. When a tax policy produces both winners and losers, there needs to be a rule according to which the government compares the utilities[9].

From the 19th century onwards utilitarianism has probably been the most common basis for social objectives. It suggests that social welfare is built up from a sum of individual utilities. To reflect the value judgements of the government this sum can be weighted. The utilitarian social welfare function can be grounded

---

[7] Admittedly, several branches of the literature have been omitted despite their importance and crucial role in the research of optimal taxation.

[8] For a discussion on the social objectives in optimal tax models, see Tuomala (1990).

[9] Measurement and comparability of individual utility is not self-evident. As not all individual characteristics are observable, utility functions may not include all the relevant information. For an extensive discussion on the problems of using utility as a criterion for policy planning, see Sen (1977, 1982)

for example by individuals' choice under risk: individuals will choose that income distribution that maximises their expected utility, the arithmetic mean of individual utility (Harsanyi, 1955).

An alternative social welfare function often mentioned is *maximin* introduced by Rawls (1971). It states that society should maximise the utility of the worst-off member and ignore all other welfare comparisons. This is rather an extreme case, but technically easy to use in theoretical models and also useful in the comparison of the results with different social objectives.

Both utilitarian and Rawlsian social welfare functions can be counted to be welfarist, i.e. depend only on the individual utilities. However, it has been argued that individuals' own assessments of their utility may not be correct, even in their own opinion. This might justify a non-welfarist, or paternalist, social welfare function, where the form of the utility used in the social welfare function differs from that used by individuals. Paternalism as a criterion for social planning has lately received a lot of attention with the emergence of behavioural economics considered in the next section.

However, accepting paternalism as a rule for social objectives is not an uncontroversial question. On the one hand, it has been shown that people actually make faulty decisions that they regret later, in which case individuals might want the government to intervene. On the other hand, it can always be argued that the sovereignty of the individual should be left untouched, and it is always possible that the government also makes similar mistakes with utility calculations.[10] In a recent discussion Thaler and Sunstein (2003) speak in favour of paternalism whereas Glaeser (2006) presents a more critical view on the issue.

[10] See LeGrand (2003) for a discussion.

In this thesis the government's objective functions are fixed[11] and all types of social welfare functions will be used: the first and second essays consider the Pareto criterion for taxation, in the third essay the government is assumed to have paternalistic preferences and in the fourth essay both utilitarian and paternalistic social welfare functions are used.

Government's benevolence and ability to commit

The assumption of benevolent government is often criticised. However, even if this assumption might not be fulfilled in real life, it is still useful to consider what policies one would want a benevolent government to follow and what the link between policies and social outcomes is. Buchanan and Musgrave (1999) compare the importance and usefulness of visualising a correctly working public sector to the consideration of *homo economicus* or competitive Walrasian system as fictions of ideal market. However, governments' actual workings should have a crucial role in the design of policy recommendations on the results received from basic theory.

One essential assumption of the optimal tax models considers the time horizon. In reality consumption choices and tax systems are naturally dynamic, in which there are several 'rounds' of agents' choices regarding consumption and labour supply and government's choice for the tax system. It has been shown that a consistent solution in general is not optimal (Kydland and Prescott, 1977, 1980): in a several-period model the government has an incentive to use the information gathered in earlier periods and deviate from the consistent tax system. This

---

[11] Instead of taking government's objectives exogenously, in the branch of political economy there is a wide body of literature considering how the political process itself affects the formation of government's objective functions. This, however, is left outside the scope of this thesis.

problem remains even if the government is benevolent (Roberts, 1984). The essential assumption is the level of commitment: if the government can be fully committed not to use the information collected in the previous periods, the solution to a dynamic model coincides with that in the static model. There is a wide body of literature considering optimal taxation with time inconsistent policies and the means to tackle with the problem.[12] However, three of the essays are employed in a static, one-period model, and one of the essays considers a two-period version of the model. Thus, the consideration of the dynamic optimal tax problem[13] is left outside the scope of this thesis.

The behavioural aspect of taxation

Behavioural economics integrates psychology with economic theory.[14] The realisation of the fact that people do not behave consistently, as would be expected from a *homo economicus*, is not a recent finding. In the time of moral philosophers human behaviour was closely linked to economics. Bernoulli (1738) already obscured that individuals are risk averse and that the level of risk aversion varies with wealth. Later, in the late 19th century in the neoclassical era economics evolved towards the natural sciences and an assumption of rational behaviour became dominant. Psychology, as a new discipline, was considered to be too unscientific and unstable as a basis for economic theory. Despite some writing by

---

[12] See for example Kehoe (1989), who suggests that tax competition might work as a partial commitment and Konrad (2001), who finds that the welfare cost accruing from the time inconsistency problem in optimal taxation is reduced by government's imperfect information of the agents' earning abilities.

[13] For a survey of optimal taxation in a dynamic setting see Golosov, Tsyvinski and Werning (2006).

[14] For a review of the emergence and development of behavioural economics, see e.g. Camerer and Loewenstein (2004)

economists such as I. Fisher, V. Pareto and J.M. Keynes on people's feelings and thoughts on economic decision-making, in the first part of 20th century the psychological insights in economics were largely omitted.

The significance of psychology in economic modelling was brought back to the agenda as models with expected utility and discounted utility were shown to include anomalies. It was shown by examples that the choices people make are not consistent with rational expectations[15]. Bounded rationality (Simon, 1955), prospect theory (Kahneman and Tversky 1979) and hyperbolic discounting (Laibson, 1997) are formal models that aim to explain the inconsistencies in people's choices under risk.

Inconsistencies in individuals' behaviour give a rationale for government intervention. When individual utility maximisation leads to an inefficient outcome there is a justification for the government to use paternalist social welfare function as a determinant for the tax policy. Widely used examples of mistakes resulting from time inconsistency are obesity and alcoholism. At the time when goods are consumed the future disutility from lost health is undervalued by individuals but the social planner might be able to assess the future disutility more accurately.[16] The analysis of such corrective public actions is called behavioural public economics.

The behavioural aspect of optimal taxation offers a new way to consider the old and new policy implications of the models. Some of the future challenges, like aging and growing health problems can be better understood with models including time inconsistencies. Thus, these characteristics should also be taken into account in the design of optimal tax policy. In this thesis, paternalism and its justification as a criterion for social welfare are discussed in the last two essays.

---

[15] Probably the most famous first examples are Allais' (1953) and Ellsberg's (1961) paradoxes.

[16] For an example of the time inconsistency problem, see O'Donoghue and Rabin (2003) considering taxing goods with delayed negative health effects, like unhealthy food or alcohol.

Multidimensional optimisation problems

Most of the research on optimal taxation has been done using models with a representative consumer or, to make redistribution a relevant question, with individuals differing in their wage earning abilities following Mirrlees (1971). In reality individuals differ in more characteristics than just that one. Moreover, it is very unlikely that all individuals have identical utility functions. Combining differences in earning abilities and in some other characteristics leads to models with multidimensional heterogeneity.

When heterogeneity emanates from preferences, comparison of the groups with different utilities is no longer clear, as was pointed out by Sandmo (1993). The fairness of redistribution needs to be carefully evaluated: is it justified to redistribute income from a person who likes to consume expensive goods and is willing to work hard for that purpose to people who hate working but are also satisfied with the low income resulting from their choices. This problem has been widely discussed in the literature of social choice.[17]

The biggest difficulty with a model with multiple unobservable characteristics is differentiating between groups. It is no longer clear which group is better off revealing their true characteristics and which group would prefer to be associated with another group. In technical terms, the pattern of binding incentive-compatibility constraints becomes unclear. Multidimensional screening problems have been discussed in the context of non-linear pricing and monopoly regulation.[18] In the framework of optimal taxation, multidimensional heterogeneity was first considered in the context of using tax policies, like

[17] For a survey, see e.g. Fleurbaey and Maniquet (1999).

[18] See e.g. Wilson (1993, 1995) Rochet and Choné (1998) and Armstrong and Rochet (1999).

workfare or indirect taxation, as a tool for screening.[19] When individuals differ with respect to more than one characteristic, it might even be desirable to choose a tax schedule that does not seem to be optimal in terms of observable characteristics (Boadway et al., 2002).

In order to obtain some analytical results, multidimensional optimisation problems are usually considered in a simple model with a discrete number of types and two-dimensional heterogeneity.[20] The discrete case is considered mainly as a three-type model to minimise the complexities.[21] However, due to the screening problem discrete models also include a lot of constraints and general analytical results are hard to get. Thus in many cases numerical solutions are provided to illustrate the results.

Even numerical approach faces severe problems as the number of constraints increases. With a continuum of the individuals' optimisation problems with multidimensional heterogeneities end up with a system of differential equations, which usually require numerical solutions with rather strong assumptions (see e.g. Mirrlees, 1986; Wilson, 1993, 1995; Tarkiainen and Tuomala, 1999; 2007). With a discrete model, too, the numerical solvability of the problem is not self-evident, as discussed in Judd and Su (2006) who have aimed to extend the optimal tax literature towards both multidimensional heterogeneity in static and in dynamic frameworks.

---

[19] The case of workfare was considered e.g. in Besley and Coate (1995) and Cuff (2000) and the use of indirect taxation e.g. in Saez (2002a), Jordahl and Micheletto (2002) and Blomquist and Christiansen (2004).

[20] Mirrlees (1986) considered the optimal tax model with m-dimensional populations ending up with a solution with a group of partial differential equations that is even numerically very challenging to solve.

[21] Two-dimensional heterogeneity of agents is considered in a 3-type model e.g. in Cuff (2000), Jordahl and Micheletto (2002) and Blomquist and Christiansen (2003, 2004).

The current research with multidimensional heterogeneity has provided new aspects for policy discussions. For example, contrary to the earlier theoretical findings, the taxation of capital income becomes desirable. The richer characterization of the agents brings economic models closer to everyday decision making. Also, the findings from the theoretical models support the policies that are in fact used in many countries. The multidimensional screening problem is considered in three essays, and in one of them the problem is also deliberated numerically.

A short introduction to the essays

This thesis endeavours to contribute to the literature of optimal tax theory by taking steps towards a more general picture of the determinants of tax policy. An effort to make more accurate assumptions certainly makes the model more plausible, but often at the cost of complicating the analysis. The beauty of a simple model lies in its ability to offer a clear intuition. This challenge is met here by gradually alleviating the assumptions. Thus, the comparability to earlier research is maintained and the intuition behind the models is easier to follow.

The essays consider different forms of tax policy, income tax, commodity tax and public provision, from a microeconomic perspective with an ideally working, benevolent government. The trade-off between efficiency and equity is taken into account in each case. As equity considerations are reasonable only in models with heterogeneous individuals, we have assumed that workers[22] differ in their wage earning ability. The same framework underlies all four essays: the optimal tax

[22] We have made a simplifying assumption of perfectly competitive labour market with all agents working. Thus we have omitted the distinction between intensive and extensive labour market decision discussed e.g. in Saez (2002b).

problem under asymmetric information introduced by Mirrlees (1971). The agents are assumed to have different wage earning abilities. Instead of a continuous model a discrete case following Stern (1982) and Stiglitz (1982) is considered. This model is generalised in several directions by alleviating the assumptions made in the earlier literature.

In addition to differences in wage earning ability, an additional source of heterogeneity is considered in the last three essays. The second dimensions for heterogeneity are: preference for leisure (essay 2), level of rationality (essay 3) and time preferences (essay 4). In the second and third essays a three-type case is considered, whereas the fourth essay also illustrates a general four-type case. The last essay also takes advantage of numerical simulations in order to struggle with the problems with multiple incentive constraints in a four-type model.

The two first essays consider taxation as a means to influence environmental quality. The first essay contemplates the problem in a general equilibrium framework. Contrary to the earlier literature the wage rates are not fixed but they are determined on the labour market. In the second essay the framework is drawn back to fixed wages, but the basic model is extended to the case with agents differing in two dimensions.

The trend of behavioural public economics is followed in the third and fourth essays. Both of them consider cases where government’s preferences differ from those of the individuals, i.e. there are paternalistic objectives. The third essay, which is co-authored with Jukka Pirttilä, considers optimality of public provision in a case where there are both redistributional and paternalistic objectives and agents differ in two dimensions. The fourth essay, which is co-authored with Matti Tuomala, studies the optimality of taxing savings in a case where some of the individuals are assumed to be too short-sighted to save adequately on a voluntary basis.

Essay 1: Optimal Tax Policy and Environmental Externality: A General Equilibrium Analysis

The growing concern about the environment has generated discussion on considering green taxes as a part of general public policy. In addition to environmental efficiency, redistributional objectives are also considered. There is empirical evidence showing that the direct impacts of environmental taxes are especially likely to be regressive.[23] Thus there may be a conflict between the aim of promoting environmental quality and the objective of more equal income redistribution.

The first essay studies the optimal tax policy in the presence of environmentally harmful consumption. The study uses a mixed taxation scheme with non-linear income taxes and linear commodity taxes in a two-type version of the Mirrlees model where individuals differ in their wage earning abilities. Contrary to the earlier literature on optimal tax policy with environmental externalities, the assumption that wages should be fixed is abandoned.

According to widely cited results on the structure of the optimal commodity tax levied on a good creating a harmful externality, Sandmo's (1975) 'additivity property' and Dixit's (1985) 'principle of targeting', the commodity tax meant to internalise the effect of a harmful externality should be imposed only on the good creating the externality and the internalizing part of taxation should appear additively in the tax rate. It is shown that this principle is more general than earlier assumed; it continues to also hold in a general equilibrium framework with endogenous factor prices.

The possible conflict between the environmental and redistributional objectives is considered with the help of the valuation of the harmfulness of the

[23] A survey on the research on the economics of environmental taxes can be found e.g. in OECD (2006).

externality. The direct harm to consumers from the externality is positive, but the indirect effect originating from the redistributional aspect in taxation decreases the valuation of the harmfulness of the externality when the utility from environmental quality increases with leisure. Thus, there is a possibility for a conflict between environmental objectives and redistributional aims.

Essay 2: Optimal Tax Policy with Environmental Externalities and Heterogeneous Preferences

The second essay continues analysing the optimal tax policy in the presence of a harmful environmental externality. The two-type Mirrlees model is extended by introducing an additional dimension of heterogeneity; instead of differing only in their wage earning abilities, agents also differ with respect to their preferences. The study uses a simpler three-type model where some of the high productivity types have a stronger preference for leisure than others.

Two-dimensional heterogeneity of agents complicates the analysis in several aspects. First, the direction of the redistribution needs to be determined; because of the heterogeneous preferences, it is no longer straightforward to determine which group should be treated the most leniently. The social welfare approach aims at taking individuals' different preferences into account even when it means treating otherwise similar individuals differently. On the other hand, redistribution from somebody with higher income due to his lower preference for leisure is also controversial. Second, there is a possibility of different types of optima; it might be the case that instead of separating optimum where each type chooses a distinct bundle of income and consumption, the economy might end up in a pooling optimum, where one or more groups of households are indistinguishable.

The valuation of the harmfulness of the externality includes some terms induced by the redistribution constraint that affects in the opposite direction from environmental aims. Although the existence of such terms does not imply that environmental and redistributional aspects are contradictory, it might be the case that part of the harm from the externality is compensated by gains in redistribution.

With two-dimensional unobservable characteristics any action revealing households' true type under asymmetric information is of special interest. In this essay it is shown that in the pooling optimum commodity tax can be used to differentiate and redistribute income between otherwise indistinguishable households, as long as they choose different amounts of consumption.

Another interesting question considered in the essay is the effect of the externality on commodity taxes. In the pooling optimum the externality based part of the commodity tax can no longer be separated from the other part of the tax but also affects the tax rate of the good not creating the harm, i.e. the Sandmo-Dixit principle considered in the first essay fails to hold. However, in the separating optimum we can generalise the Sandmo-Dixit principle to a model with two-dimensional heterogeneity of agents.

Essay 3: Pawns and Queens Revisited: Public Provision of Private Goods When Individuals Make Mistakes[24]

Recent findings in behavioural economics have demonstrated that individual decision-making suffers from bounded rationality and various biases. There are a number of reasons why individuals are particularly prone to make mistakes, for instance, in decisions related to health; the information required for rational decision-making may be too great or complicated, or individuals may undervalue the returns on health investment accruing later in the future. In these situations, treating the customers of public services as 'pawns', whose decisions are mainly

[24] The essay is co-authored by Jukka Pirttilä, and published in *International Tax and Public Finance*.

delegated to the provider, instead of 'queens', sovereign consumers, can therefore be desirable to induce behaviour that is closer to what individual wish they were doing.

One important area where the government could improve upon individual choice is related to goods such as education, health and insurance, which in many countries are indeed often publicly regulated, provided or subsidised. While price subsidies and public provision can both be used as corrective devices, public provision can be especially important in situations with individual mistakes. By public provision, the government can make sure that the individuals consume at least a certain minimal amount of goods deemed meritorious.

In the third essay we consider optimal tax policy with redistributive objectives and the public provision of a good that is undervalued by individuals relative to government's view. We examine what will happen to the optimal policy if the government is paternalistic, i.e. tries to correct these mistakes by basing its own decision on what it thinks is truly best for the individuals. We account for two potential constraints to optimality: asymmetric information of agents' productivity and mistakes in individual decision-making. Intuitively, individuals' undervaluation which is to be rectified should imply desirability of public provision. According to the earlier literature, public provision can be a useful tool for redistribution when it helps to reduce the harmful incentive effects of income taxation. It is therefore interesting to examine whether the two motivations for public intervention interact when individuals differ in two dimensions.

In the third essay it is shown that when all individuals are irrational and they differ only intheir productivity, public provision of the undervalued good is welfare improving. When we extend the model so that there are also differences in rationality, we can consider the interaction of the two motivations for public intervention. It is found that even if both the aim to reduce the harm from irrationality and asymmetric information separately would speak for a positive level of public provision, when these two aims are put together, it is no longer clear that public provision improves welfare. It is also found that income tax rules are affected by the connection between the undervalued good and labour supply.

Essay 4: On Optimal Lifetime Redistribution Policy[25]

Publicly provided retirement programmes can be justified on several grounds, like market failures generated by asymmetric information, redistributive grounds or myopic behaviour, i.e. some individuals might consume “excessively” during their earning years finding themselves with insufficient saving in retirement. In the situations where there is a possible conflict between individual’s preference for the long run and his or her short run behaviour, a government intervention may be desirable.

The notion that individuals may not make the best choices for themselves raises difficult issues: it is no longer clear whether the government should maximise individual welfare as the individual sees it, or, as suggested in recent behavioural public economics literature, if it should be paternalistic and discount the future at a lower rate than individuals. Individuals may be fully rational and they just happen to have a high preference for the present, which causes them to save little, because too little thought is given to future contingencies.

In the fourth essay we consider a two-period variant of the Mirrlees (1971) income tax problem, where individuals work and then retire. We characterise optimal redistribution policy within a cohort when society consists of individuals who do not differ only in productivity, but also in time preference or myopia. We solved the optimal tax treatment for savings both with welfarist and paternalistic government objectives.

[25] The essay is co-authored by Matti Tuomala.

In addition to a similar three-type model as considered in the second and third essays, here we also solve a four-type model. The problem with multiple constraints is solved with the help of numerical simulation. A numerical solution also enables us to consider replacement rates, their dependence on the individuals' characteristics and consumption dispersion in both periods. When insufficient saving is caused by myopia or low discount factor, our analytical and numerical results support the view that there is a case for a non-linear public pension programme in a world in which individuals differ in skills and discount factor or myopia.

Conclusions

This thesis consists of four independent essays all considering optimal tax policy. The same basic model introduced by James Mirrlees (1971) is used as a framework in all these essays. The model offers technical tools to consider optimal taxation and the trade-off between equity and efficiency, i.e. a framework that also takes redistributive objectives into account.

The main objective of this thesis is to extend the theoretical framework behind the optimal tax results to correspond better with the real characteristics and behaviour of the agents. The assumptions are alleviated gradually to maintain the tractability and comparability of the results. Models closer to real life also improve the policy relevance of optimal tax theory. In general, the extended models show that the basic rules from the optimal tax literature do not necessarily hold with more complex assumptions. In fact, the results can be reversed.

Although the assumptions behind the models considered in this thesis are relaxed to be closer to the real world, they still fall short in describing the actual circumstances faced in everyday decision-making. However, the extensions presented in this thesis are steps towards models closer to real world phenomena. The possible directions in which to extend optimal tax theory include. dynamic models with overlapping generations, general equilibrium models and models with a continuum of agents differing in several characteristics. As the complexity of the models starts to restrict the usefulness of analytical tools, both numerical approach and empirical consideration would offer illuminating aspects for the analysis.

References

Allais, M. (1953), ‘Le Comportements de l’Homme Rationnel devant le Risque, Critique des Postulats et Axiomes de l’Ecole Américaine’, *Econometrica* 21, 503-546.

Armstrong, M. and J.-C. Rochet (1999), ‘Multi-Dimensional Screening: - A User’s Guide’, *European Economic Review* 43, 959-979.

Aronsson, T. and T. Sjögren (2003), ‘Income taxation, commodity taxation and provisi9on of public goods under labor market imperfections’, *FinanzArchiv* 59, 347-370.

Aronsson, T. and T. Sjögren (2004), ‘Is the Optimal Labor Income Tax Progressive in a Unionized Economy?’, *Scandinavian Journal of Economics* 106, 661-675.

Atkinson, A. B. (1973), ‘*How Progressive Should Income Tax Be?*’ In J. M. Parkin and A. Nobay (eds.) ‘*Essays in Modern Economics*’, Macmillan, New York.

Atkinson, A. B. and A. Sandmo (1980), ‘Welfare Implications of the Taxation of Savings’, *Economic Journal* 90, 529-549.

Atkinson, A. B. and J. E. Stiglitz (1972), ‘The Structure of Indirect Taxation and Economic Efficiency’, *Journal of Public Economics* 1, 97-119.

Atkinson, A. B. and J. E. Stiglitz (1976), ‘The Design of Tax Structure: Direct versus Indirect Taxation’, *Journal of Public Economics* 10, 55-75.

Auerbach, A. J. and J. R. Hines (2001), ‘Perfect Taxation with Imperfect Competition’, *NBER Working Paper* No 8138.

Bernoulli, D. (1738), ‘Specimen Theoriae Novae de Mensura Sortis’ *Commentarii Academiae Scientiarum Imperialis Petropolitanae,* tonus V (*Papers of the Imperial Academy of Sciences in St. Petersburg, vol V*). Translated in D. Bernoulli (1954) ‘Exposition of a New Theory on the Measurement of Risk’, *Econometrica* 22, 23-36.

Besley, T. (1988), 'A Simple Model for Merit Good Arguments', *Journal of Public Economics* 35, 371-384.

Besley, T. and S. Coate (1995), ‘The Design of Income Maintenance Programmes’, *Review of Economic Studies* 62, 187-221.

Bergson, A. (1938), ‘A Reformulation of Certain Aspects of Welfare Economics’, *Quarterly Journal of Economics* 52, 310-334.

Blomquist, S. and V. Christiansen (2003), ‘Is There a Case for Public Provision of Private Goods if Preferences are Heterogeneous? An Example with Day Care’, CESifo Working Paper No. 938.

Blomquist, S. and V. Christiansen (2004), ‘Taxation and Heterogeneous Preferences’ CESifo Working Paper No.1244.

Boadway, R., M. Marchand, P. Pestieau and M. del Mar Racionero (2002), ‘Optimal Redistribution with Heterogeneous Preferences for Leisure’, *Journal of Public Economic Theory* 4, 475-498.

Boiteaux, M. (1956), ‘Sur la Gestion des Monopoles Publics Astreints à l’Équilibre Budgétaire’, *Econometrica* 24, 22-40. Translated in M. Boiteaux

(1971) ‘On the Management of Public Monopolies Subject to Budgetary Constraints’, *Journal of Economic Theory* 3, 219-240.

Buchanan, J. M. and R. A. Musgrave (1999), ‘*Public Finance and Public Choice*’, MIT Press, Cambridge.

Camerer, C. F and G. Loewenstein (2004), ‘*Behavioural Economics: Past, Present, Future*’ in C. F. Camerer, G. Loewenstein and M. Rabin (eds.) (2004), ‘*Advances in Behavioural Economics*’, Russel Sage Foundation, New York.

Cassel, G. (1901), ‘The Theory of Progressive Taxation’, *Economic Journal* 11, 481-491.

Corlett, W. J. and D. C. Hague (1953), ‘Complementarity and the Excess Burden of Taxation’, *Review of Economic Studies* 21, 21-30.

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

Deaton, A. S. (1977), ‘Equity, Efficiency and the Structure of Indirect Taxation’, *Journal of Public Economics* 8, 299-312.

Diamond, P. A. (1975), ‘A Many-Person Ramsey Tax Rule’, *Journal of Public Economics* 4, 335-346.

Diamond, P.A. and J. Mirrlees (1971), ‘Optimal Taxation and Public Production I: Production Efficiency and II: Tax Rules’, *American Economic Review* 61: 8-27 and 261-278.

Dixit, A. (1985), ‘A Tax Policy in Open Economies’, in A.J. Auerbach and M. Feldstein (eds.), ‘*Handbook of Public Economics, vol. 1*’, North-Holland.

Dixit, A. and A. Sandmo (1977), 'Some Simplified Formulae for Optimal Income Taxation', *Scandinavian Journal of Economics* 79, 417-423.

Ellsberg, D. (1961), 'Risk, Ambiguity, and the Savage Axioms', *Quarterly Journal of Economics* 75, 642-669.

Edgeworth, F. Y. (1897), 'The Pure Theory of Taxation', *Economic Journal* 7, 46-70, 226-238, 550-571. Reprinted in F.Y. Edgeworth (1925), '*Papers Relating to Political Economy*', Macmillan, London.

Edgeworth, F. Y. (1919), 'Methods of Graduating Taxes on Income and Capital', *Economic Journal* 29, 144-165. Reprinted in F.Y. Edgeworth (1925), '*Papers Relating to Political Economy*', Macmillan, London.

Fair, R. C. (1971), 'The Optimal Distribution of Income', *Quarterly Journal of Economics* 85, 551-579.

Feldstein, M. (1985), 'The Optimal Level of Social Security Benefits', *Quarterly Journal of Economics* 100, 300-320

Fleurbaey, M. and F. Maniquet (1999), 'Compensation and Responsibility', Mimeo, revised in 2003. Forthcoming in Arrow, K. J., A. K. Sen and K. Suzumura (eds.) '*Handbook of Social Choice and Welfare, Vol 2*', North-Holland.

Fuest, C. and B. Huber (1997), 'Wage bargaining, labor-tax progression, and welfare", *Journal of Economics* 66, 127-150.

Glaeser, E. L. (2006), 'Paternalism and Psychology', *The University of Chigago Law Review* 73, 133-156.

Golosov, M., A. Tsyvinski and I. Werning (2006), 'New Dynamic Public Finance', *NBER Macroeconomics Annual* 2006.

Graaff, J. de van (1957), ‘*Theoretical Welfare Economics*’, Cambridge University Press.

Guesnerie, R. and J. Seade (1982), ‘Nonlinear Pricing in a Finite Economy’, *Journal of Public Economics* 17, 157-179.

Harsanyi, J. (1955), ‘Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility’, *Journal of Political Economy* 63, 309-321.

Hudson, M. (2000), ‘Land-Value Taxation Around the World: Mesopotamia and Classical Antiquity’, *The American Journal of Economics and Sociology* 59 (5), 3-25.

Jordahl, H. and L. Micheletto (2002), ‘Optimal Utilitarian Taxation and Horizontal Equity’, Working Paper Series 2002:19, Uppsala University, Department of Economics.

Judd, K. L. and C.-L. Su (2006), ‘Optimal Income Taxation with Multidimensional Taxpayers Types’, Computing in Economics and Finance 2006, 471, Society for Computational Economics.

Kahneman, D. and A. Tversky (1979), ‘Prospect Theory: An Analysis of Decision Under Risk’, *Econometrica* 47, 263-281.

Kehoe, P. (1989), ‘Policy Cooperation among Benevolent Governments May Be Undesirable’, *Review of Economic Studies* 56, 289-296.

Konrad, K. (2001), ‘Privacy and Time-Consistent Optimal Labor Income Taxation’, *Journal of Public Economics* 79, 503-519.

Kydland, F. and E. C. Prescott (1977), ‘Rules Rather than Discretion: The Inconsistency of Optimal Plans’, *Journal of Political Economy* 85, 473-492.

Kydland, F. and E. C. Prescott (1980), 'Dynamic Optimal Taxation, Rational Expectations and Optimal Control', *Journal of Economic Dynamics and Control* 2, 79-91.

Laibson, D. (1997), 'Golden Eggs and Hyperbolic Discounting', *Quarterly Journal of Economics* 116, 443-477.

LeGrand, J. (2003), '*Motivation, Agency, and Public Policy. Of Knights & Knaves, Pawns & Queens*', University Press, Oxford.

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

Mirrlees, J. A. (1976), 'Optimal Tax Theory: A Synthesis', *Journal of Public Economics* 20, 77-87.

Mirrlees, J. A. (1986), 'The Theory of Optimal Taxation' in K. J. Arrow, and M. D. Intriligator (eds.) in '*Handbook of Mathematical Economics, vol. III*', Elsevier, Amsterdam.

Musgrave, R. A. (1959), '*The Theory of Public Finance*', McGraw-Hill.

Musgrave, R. A. (1985), '*A Brief History of Fiscal Doctrine*' in A. J. Auerbach and M. Feldstein (eds.) '*Handbook of Public Economics, vol 1*', North-Holland.

Nichols, A. and R. Zeckhauser (1982), 'Targeting Transfers through Restrictions on Recipients', *American Economic Review, Papers and Proceedings* 72, 372-377.

OECD (2006), '*The Political Economy of Environmentally Related Taxes*', OECD.

O'Donoghue, T. and M. Rabin (2003), 'Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes', *American Economic Review* 93, 186-191.

Ordover, J. A. and E. S. Phelps (1979), 'The concept of Optimal Taxation in an over-lapping generations model of capital and wealth', *Journal of Public Economics* 12, 1-26.

Pigou, A. C. (1928), '*A Study in Public Finance*', MacMillan.

Pigou, A. C. (1947), '*A Study in Public Finance,* 3rd Edition', Macmillan, London.

Ramsey, F. P. (1927), 'A Contribution to the Theory of Taxation', *Economic Journal* 38, 543-559.

Rawls, J. (1971), '*A Theory of Justice*', Harvard University Press, Cambridge.

Robbins, L. (1932), '*An Essay on the Nature and Significance of Economic Science*', MacMillan.

Robbins, L. (1938), 'Interpersonal Comparison of Utility: A Comment', *Economic Journal* 48, 435-641.

Roberts, R. (1984), 'A Positive Model of Private Charity and Public Transfers', *Journal of Political Economy* 92, 136-148.

Rochet, J.-C. and P. Choné (1998), 'Ironing, Sweeping and Multidimensional Screening', *Econometrica* 66, 783-826.

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

Saez, E. (2002b), ‘Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses’, *Quarterly Journal of Economics* 117, 1039-1073.

Salanié, B. (2003), ‘The Economics of Taxation’, MIT Press, Cambridge.

Sandmo, A. (1975), ‘Optimal Taxation in the Presence of Externalities’, *Swedish Journal of Economics* 77, 86-98.

Sandmo, A. (1993), ‘Optimal Redistribution When Tastes Differ’, *FinanzArchiv* 50(2), 149-163.

Schumpeter, J. A. (1954), ‘*History of Economic Analysis*’, Oxford University Press.

Screpanti, E. and S. Zamagni (2005), ‘*An Outline of the History of Economic Thought*’, Oxford University Press.

Sen, A. (1977), ‘On Weights and Measures: Informational Constraints in Social Welfare Analysis’, *Journal of Public Economics* 3, 387-403.

Sen, A. (1982), ‘*Choice, Welfare and Measurement*’, Blackwell, Oxford.

Sheshinski, E. (1972), ‘The Optimal Linear Income Tax’, *Review of Economic Studies* 39, 297-302.

Sidgwick, H. (1883), ‘*Principles of Political Economy*’, Macmillan, London.

Simon, H. A. (1955), ‘A Behavioral Model of Rational Choice’, *Quarterly Journal of Economics* (69), 99-118

Stern, N. H. (1976), ‘In the Specification of Optimum Income Taxation’, *Journal of Public Economics* 6, 123-162.

Stern, N. H. (1982), ‘Optimum 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.

Stiglitz, J. E. and P. S. Dasgupta (1971), ‘Differential Taxation, Public Goods and Economic Efficiency’, *Review of Economic Studies* 38, 151-174.

Thaler, R. H. and C. R. Sunstein (2003), ‘Libertarian Paternalism’, *American Economic Review* 93, issue 2, 175-179.

Tuomala, M. (1990), ‘*Optimal Income Tax and Redistribution*’, Clarendon Press, Oxford.

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

Tarkiainen, R. and M. Tuomala (1999), ’On Optimal Income Taxation with Heterogeneous Work Preferences’, *International Journal of Economic Theory* 3, 35-56.

Vickrey, W. S. (1945), ‘Measuring Marginal Utility by Reactions to Risk’, *Econometrica* 13, 319-333.

Wesson, J. (1972), ‘On the Distribution of Personal Incomes’, *Review of Economic Studies* 39, 77-86.

Wilson, R. (1993), ‘*Nonlinear Pricing*’, Oxford University Press, New York.

Wilson, R. (1995), '*Nonlinear Pricing and Mechanism Design*' in H. Amman, D. Kendrick and J. Rust (eds.), '*Handbook of Computational Economics, Vol 1.*', Elsevier, Amsterdam.

Zeckhauser, R. (1970), '*Uncertainty and the Need for Collective Action*' in R. Haveman and J. Margolis (eds.) '*Public Expenditure and Policy Analysis*', Markham, Chigago.

Chapter 2

Optimal Tax Policy and Environmental Externality: A General Equilibrium Analysis

Sanna Tenhunen[♣]
University of Tampere & FDPE

Abstract: This study analyses optimal tax policy in a two-type world with environmental externalities, endogenous wages and mixed taxation scheme. The possible contradiction in combining environmental and redistributional objectives is studied via terms in the valuation of the externality. The presence of the externality tends to increase the marginal effective tax rates of both types, whereas the effect of endogeneous wages decreases the marginal effective tax rate of the high productivity type and increases that of the low productivity type. However, the Sandmo-Dixit principle continues to hold in commodity taxation: the externality internalizing part of the commodity tax appears additively only in the tax rate of the dirty good.

Keywords: Optimal taxation, externality, Sandmo-Dixit principle
JEL: D82, H21, H23

[♣] Address: Department of Economics and Accounting, FIN-33014 University of Tampere.
E-mail: sanna.tenhunen@uta.fi.

1. Introduction

Taxation is a commonly accepted instrument in encouraging the economy to limit environmentally damaging activities. Imposing a Pigouvian tax on such a polluting good internalises the externality, and a distorting tax is therefore beneficial to the economy from the efficiency point of view. In addition to efficiency, redistributional objectives are also an important aspect of optimal tax policy. Diamond and Mirrlees (1971) introduced a theory of optimal commodity taxation where both these aspects were concerned. The first study combining environmental issues with the redistribution aspect of taxation is Sandmo (1975), where the structure of commodity taxation is analysed.

There is some empirical evidence showing that environmental taxes are regressive.[1] As governments usually have some kind of redistributional goal, there may be a conflict between the aims of promoting both environmental quality and income redistribution. In this study we try to find evidence from theoretical point of view on the question as to whether environmental and redistributive objectives are contradictory.

The aim of this study is to analyse optimal tax policy in the presence of an environmental externality. Our study uses a mixed taxation scheme with non-linear income taxes and linear commodity taxes introduced by Mirrlees (1976) and Atkinson and Stiglitz (1976). The model includes two types of households differing in their earning abilities, following the path of Stiglitz (1982) and Stern (1982). This division of the households in the model also enables the consideration of redistributional issues. The true type of the household cannot be detected; only the income level is observable. As a result of this asymmetric

[1] According to most of the literature, the direct impacts of environmental taxes are especially likely to be regressive. See e.g. Smith (1992), Harrison (1994) and for a more recent study on energy taxes Brännlund and Nordström (2004) and West and Williams (2004).

information the self-selection constraint needs to be taken into account when designing optimal tax policy. A model combining the assumptions of a two-type economy and mixed taxation is used e.g. in Edwards, Keen and Tuomala (1994). They find that commodity taxes can be used to weaken the self-selection constraint. Pirttilä and Tuomala (1997) develop the model further; their analysis of optimal tax policy includes a harmful externality affecting environmental quality.

In most analyses the factor prices are assumed to be exogenous. Naito (1999) used endogenous factor prices in the two-type economy to show that factor price determination has essential effects on taxation. Aronsson and Sjögren (2003) consider optimal taxation in a case where labour markets are not perfect, but wages are instead determined by unions. Micheletto (2004) studied also the effects of factor price endogeneity on optimal tax policy. Our aim is to combine the assumptions of environmental externalities and endogenous wages into a two-type economy with mixed taxation.

The possible conflict between environmental and redistributional objectives is considered with the help of the valuation of the harmfulness of the externality. Consumer side terms in the valuation of the externality turn out to have opposite signs; the direct harm to consumers is positive, but the indirect effect originating from the self-selection constraint decreases the valuation of the externality when environmental quality and leisure are complements. Smaller income differences decrease the effect of self-selection constraint by cutting the incentives to mimic and thus increase the harmfulness of the externality. On the producer side the direct effect in production and the indirect effect of the wage adjustment on the labour markets both increase the valuation of the externality. Thus on the consumer side there is a possibility for contradiction between environmental and redistributional objectives, whereas on the producer side environmental targets are in accordance with the income distribution objectives.

The paper is constructed as follows. Section 2 introduces the model, maximisation problem and the first order conditions defining the optimal tax policy. In Section 3 we derive the valuation of the externality and interpretations for the terms in it. The optimal tax policy, i.e. the marginal effective tax rates for both types of households and the optimal commodity tax rates, are determined in Section 4. Finally, Section 5 concludes.

2. The model

2.1 Structure of the economy

There are assumed to be two types of households with similar preferences and endowments but with different productivities. Type 1 households have lower productivity than type 2 households. The productivity of the household is not observable. Without loss of generality the number of both households is normalized to one. The labour income earned by households equals $Y^h = w^h L^h$, where L^h denotes the labour supply and w^h denotes the wage rate of type *h*, *h=1,2*.

The role of the public sector is to design an optimal tax structure so that its two points of interest, environmental quality and redistribution are taken into account. It has two tax devices at its disposal: non-linear income tax and linear commodity tax. Households' labour income is taxed by a non-linear income tax scheme $T(Y^h)$ such that the after-tax income used for consumption is $B^h = Y^h - T(Y^h)$, $h =$ 1, 2. All of the tax revenue is used to finance an exogenous revenue requirement $\overline{G}$.

There are two goods X_i, $i=c,d$, denoted from now on by a vector $\mathbf{X} = \begin{bmatrix} X_c \\ X_d \end{bmatrix}$.

Good X_c is a "clean" good, whereas good X_d is a "dirty" good; the externality E damaging the environment is assumed to be created by aggregate consumption of the dirty good, $E = \sum_h X_d^h$, h =1, 2. Both goods are assumed to be normal. The goods are produced with a joint, constant returns to scale production function $F(L^1, L^2, E)$, which is increasing in both types' labour supply, i.e. $\frac{\partial F}{\partial L^h} = w^h > 0$, h = 1,2, and decreasing in the level of the externality, $\frac{\partial F}{\partial E} < 0$.

Contrary to Pirttilä and Tuomala (1997), we assume that while prices remain fixed, wages are no longer exogenously determined by the productivities. Instead, wages are assumed to be determined endogenously on the labour market, i.e. the wage rate depends on the marginal cost of production, which is not a constant.[2] Without a loss of generality it is assumed that $w^2 > w^1$. The wage ratio $\Omega = \frac{w^1}{w^2}$ now depends on labour supply and externality: $\Omega = \Omega(L^1, L^2, E)$. For the wage formation process we assume that (i) both labour types are used in the equilibrium and they are substitutes for each other in the same proportion in the production of both goods, and (ii) the externality reduces both types' productivity differently, i.e. $\frac{\partial \Omega}{\partial E} \neq 0$.

The exogenous consumer prices are denoted by a vector $\mathbf{q} = \begin{bmatrix} q_c \\ q_d \end{bmatrix}$ as a sum of producer prices $\mathbf{p} = \begin{bmatrix} p_c \\ p_d \end{bmatrix}$ and commodity taxes $\mathbf{t} = \begin{bmatrix} t_c \\ t_d \end{bmatrix}$: $\mathbf{q} = \mathbf{p} + \mathbf{t}$. The demand of household h for commodity i, X_i^h, depends on the price vector **q,** the after-tax income B^h and the labour supply L^h and the level of the externality E.

[2] For example, a small open economy where commodities are traded internationally but labour is perfectly immobile could explain this kind of a structure of prices and wages.

2.2 The optimisation problem

Households maximize their direct utility $U^h(\mathbf{X}^h, L^h, E)$ subject to budget constraint $\mathbf{q}^T\mathbf{X}^h = B^h$. This results in an indirect utility function denoted by $V^h(\mathbf{q}, B^h, L^h, E)$. We focus on the characterization of Pareto-efficient tax policies assuming that government maximises low ability type's indirect utility V^1 for a given the level of high ability type's utility V^2 (Pareto constraint).

As government cannot observe the ability of a worker, the income tax is based on the observable labour income. Thus government is restricted by a self-selection constraint, i.e. taxation needs to be defined in such a way that both household types must prefer to select the combination of labour supply and income taxation meant for them instead of mimicking the choice of the other type. Here we concentrate on the situation in which the only binding self-selection constraint is the one requiring the high ability households not to mimic the low ability ones. The income of the mimicker (referred to with a 'hat') is $\hat{Y}^2 = Y^1 = w^2\hat{L}^2$, and the labour supply for mimicker is $\hat{L}^2 = \frac{Y^1}{w^2} = \frac{w^1L^1}{w^2} = \Omega L^1$. Now the self selection constraint for the high ability type can be written as

$$V^2(\mathbf{q}, B^2, L^2, E) \geq V^2(\mathbf{q}, B^1, \Omega L^1, E). \tag{1}$$

The assumption that all tax revenues are used to finance the exogenous revenue requirement gives us the public sector's budget constraint $\sum_h (T(Y^h) + \mathbf{t}^T\mathbf{X}^h) = \overline{G}$, where capital T in $\mathbf{t}^\mathrm{T}$ denotes a transposed vector. It can be rewritten with the help of consumer's budget constraint as

$$F(L^1, L^2, E) - \sum_h \mathbf{p}^T\mathbf{X}^h(\mathbf{q}, B^h, L^h, E) = \overline{G} \tag{2}$$

The Lagrange function of government's optimisation problem is

$$
\begin{aligned}
\Psi = & V^1\left(\mathbf{q}, B^1, L^1, E\right) \\
& + \delta\left[V^2\left(\mathbf{q}, B^2, L^2, E\right) - \overline{V}^2\right] + \lambda\left[V^2\left(\mathbf{q}, B^2, L^2, E\right) - V^2\left(\mathbf{q}, B^1, \Omega L^1, E\right)\right] \\
& + \gamma\left[F\left(L^1, L^2, E\right) - \sum_h \mathbf{p}^T \mathbf{X}^h\left(\mathbf{q}, B^h, L^h, E\right) - \overline{G}\right] \\
& + \mu\left[E - \sum_h X_d^h\left(\mathbf{q}, B^h, L^h, E\right)\right]
\end{aligned} \tag{3}
$$

where δ, λ, γ and μ are the Lagrange multipliers for the Pareto constraint, the self-selection constraint for high ability type, the government's budget constraint and the constraint defining externality respectively.

The planner maximizes this function with respect to L^h and B^h, $h=1,2$, prices $\mathbf{q}$ and E. The first order conditions with respect to these variables are:

$$L^1: \quad V_L^1 - \lambda \hat{V}_L^2\left(\Omega + L^1 \frac{\partial \Omega}{\partial L^1}\right) + \gamma\left(w^1 - \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial L^1}\right) - \mu \frac{\partial X_d^1}{\partial L^1} = 0 \tag{4}$$

$$B^1: \quad V_B^1 - \lambda \hat{V}_B^2 - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial B^1} - \mu \frac{\partial X_d^1}{\partial B^1} = 0 \tag{5}$$

$$L^2: \quad (\delta + \lambda) V_L^2 - \lambda \hat{V}_L^2 L^1 \frac{\partial \Omega}{\partial L^2} + \gamma\left(w^2 - \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial L^2}\right) - \mu \frac{\partial X_d^2}{\partial L^2} = 0 \tag{6}$$

$$B^2: \quad (\delta + \lambda) V_B^2 - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2} - \mu \frac{\partial X_d^2}{\partial B^2} = 0 \tag{7}$$

$$\mathbf{q}: \quad V_{\mathbf{q}}^1 + (\delta + \lambda) V_{\mathbf{q}}^2 - \lambda \hat{V}_{\mathbf{q}}^2 - \gamma \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} - \mu \sum_h \frac{\partial X_d^h}{\partial \mathbf{q}} = \mathbf{0} \tag{8}$$

$$
E: \quad
\begin{aligned}
& V_E^1 + (\delta + \lambda) V_E^2 - \lambda\left(\hat{V}_E^2 + \hat{V}_L^2 L^1 \frac{\partial \Omega}{\partial E}\right) \\
& \qquad + \gamma\left(F_E - \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E}\right) + \mu\left(1 - \sum_h \frac{\partial X_d^h}{\partial E}\right) = 0.
\end{aligned} \tag{9}
$$

Here the subscripts refer to partial derivatives with respect to the corresponding variable, and the hat terms $\hat{V}^2$ refer to the mimicker's utility.

3. The harmfulness of externality

3.1 Derivation of the shadow price of the externality

The interpretation for a Lagrange multiplier is the shadow price for the corresponding constraint. Thus μ is the shadow price of the externality which can be interpreted as the valuation of the harmfulness of the externality. We use here the form $\frac{\mu}{\gamma}$, where the valuation of the externality is given in terms of government tax revenues.

To find an expression for the shadow price, we manipulate the first order condition in Eq. (9). A more detailed derivation is presented in Appendix A. We let $MWP_{EB}^h = -\frac{V_E^h}{V_B^h}$ denote the marginal willingness to pay to avoid the externality and we use first order conditions (5) and (7) to obtain

$$
\begin{aligned}
&-\left(\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial B^1} + \mu \frac{\partial X_d^1}{\partial B^1}\right) MWP_{EB}^1 - \left(\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2} + \mu \frac{\partial X_d^2}{\partial B^2}\right) MWP_{EB}^2 \\
&\qquad + \lambda \hat{V}_B^2 \left(M\hat{W}P_{EB}^2 - MWP_{EB}^1\right) - \lambda \hat{V}_L^2 L^1 \frac{\partial \Omega}{\partial E} \qquad (10) \\
&\qquad\qquad + \gamma \left(F_E - \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E}\right) + \mu \left(1 - \sum_h \frac{\partial X_d^h}{\partial E}\right) = 0.
\end{aligned}
$$

It is useful to define a compensated conditional demand function x_i^h obtained by minimising consumer's expenditures $\mathbf{q}^T \mathbf{X}$ subject to keeping the direct utility $U^h\left(\mathbf{X}^h, L^h, E\right)$ at a constant level u^h. The compensated conditional demand function $x_i^h\left(\mathbf{q}, L^h, E\right)$ satisfies the properties $\frac{\partial X_i^h}{\partial E} = \frac{\partial x_i^h}{\partial E} - MWP_{EB}^h \frac{\partial X_i^h}{\partial B^h}$ and

$\mathbf{p}^T \frac{\partial \mathbf{x}^h}{\partial E} = MWP_{EB}^h - \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E}$.[3] Substituting these into Eq. (10) and using a definition $\lambda^* = \frac{\lambda \hat{V}_B^2}{\gamma}$ the shadow price of the externality can be written as

$$\frac{\mu}{\gamma} = \sigma \left[\underbrace{\sum_h MWP_{EB}^h}_{\text{term } C_d} - \underbrace{\lambda^* \left(M\hat{W}P_{EB}^2 - MWP_{EB}^1 \right)}_{\text{term } C_i} - \underbrace{F_E}_{\text{term } P_d} + \underbrace{\lambda^* \frac{\hat{V}_L^2}{\hat{V}_B^2} L^1 \frac{\partial \Omega}{\partial E}}_{\text{term } LM} - \underbrace{\sum_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E}}_{\text{term } G_i} \right],$$

$$\text{where } \sigma = \frac{1}{1 - \sum_h \frac{\partial x_d^h}{\partial E}}. \tag{11}$$

The valuation of the externality consists of six terms. The coefficient σ is the environmental feedback parameter. It has been shown that the feedback parameter must be positive to ensure the stability of the model (Sandmo, 1980). The terms in brackets are divided into terms concerning consumers (C), producers (P), the labour market (LM) and the government (G) and the subscripts refer to direct effects (d) and indirect effects (i).

3.2 Interpretations for the terms in the shadow price

The first term in the brackets, C_d, describes the direct harmful effect of the externality on consumers. It is a sum of the marginal willingnesses to pay to avoid the externality. Resulting from its definition, the sum is positive as the externality is assumed to affect the utility of the households negatively.

The shadow price is affected by the self-selection constraint through the indirect effect of the externality on consumers, term C_i. Its sign depends on the

[3] See Edwards, Keen and Tuomala (1994) for a proof.

difference in valuation of the externality between mimicker and true type 1 household. Since the mimicker is of the high ability type, he can do the work of type 1 in a shorter time than a true low ability type worker, and thus he has more leisure. When environmental quality and leisure are complements (substitutes), the marginal willingness to pay is the greater the more (less) leisure household has. Thus, it can be concluded that when environmental quality and leisure are complements (substitutes), MWP_{EB} is larger for the mimicker than for the low productivity household and the term has a negative (positive) effect on the shadow price of the externality.

When environmental quality and leisure are complements, mimickers are willing to pay more to avoid the externality than true low ability type workers. Thus lowering the level of the externality raises the desirability of mimicking. In order to reduce the incentive to mimic back to the level where mimicking does not give higher utility for the high ability type than revealing their true type, government needs to increase the taxation of the low ability households. This in turn changes the level of the income redistribution towards wider income differences, which contradicts government's redistribution objective. Thus income distribution target and the objectives to promote environmental quality are not in accordance with respect to these two consumer side terms[4]. The case where environmental quality and leisure are substitutes proceeds analogously. In this case promoting environmental quality is in accordance with redistributional aims, as decreasing the level of externality makes mimicking less desirable and thus, by mitigating the self-selection constraint, allows a more equal income distribution.

---

[4] On the consumer side there are actually two indirect effects: impacts of the self-selection constraint and the labour market influencing through wage adjustment. In the next section it is shown that the sign of the labour market effect is most plausibly positive, so these two effects have opposite signs and as a whole the impact of the indirect effect on the shadow price on the consumer side remains ambiguous.

The third term inside the brackets in Eq. (11), P_d, is the marginal cost of the externality in production. Thus this term can be interpreted to be the direct effect on the producers. P_d increases the shadow price as the derivative of the production function with respect to the externality was assumed to be negative. The indirect effect on the producer side arises from labour markets. As wages are not constant, their adjustment affects the valuation of the externality. This effect appears in the second last term in Eq. (11), called the labour market term *LM*, due to the fact that the wage adjustment affects both consumers and producers.

The sign of the *LM* term depends on the partial derivative of the wage ratio Ω with respect to the externality as λ^*, L^1 and $\hat{V}_B^2$ are positive and $\hat{V}_L^2$ is negative. Since the wage ratio Ω can be defined with the help of marginal productivities, it can be written as $\Omega = \frac{w^1}{w^2} = \frac{\partial F(L^1, L^2, E)/\partial L^1}{\partial F(L^1, L^2, E)/\partial L^2} = \frac{F_1(L^1, L^2, E)}{F_2(L^1, L^2, E)}$. Thus the partial derivative Ω_E is

$$\frac{\partial \Omega}{\partial E} = \Omega\left(\frac{F_{1E}}{w^1} - \frac{F_{2E}}{w^2}\right) = \Omega \frac{1}{E}(\xi_{1E} - \xi_{2E}), \tag{12}$$

where the subscripts refer to the partial derivatives. Now the sign of Ω_E is determined by ξ_{1E} and ξ_{2E}, the elasticities of wages with respect to the externality. Assuming pollution to have a negative effect on productivity, these elasticities are negative. Ω_E is negative when $|\xi_{1E}| > |\xi_{2E}|$, i.e. when the low ability worker's productivity increases relatively more than the high ability type's productivity as the level of pollution is decreased[5]. Based on the general conclusion of empirical

[5] The reasoning is the same in the case where the effect of the externality on productivity is positive. Ω_E is positive, if a type 2 worker suffers more from the increased amount of the externality. The other possibilities are that (i) $\xi_{1E} < 0$ and $\xi_{2E} > 0 \rightarrow \Omega_E$ always negative, and (ii) $\xi_{1E} > 0$ and $\xi_{2E} < 0 \rightarrow \Omega_E$ always positive.

findings in the literature[6], we assume here that the low ability workers are more vulnerable to pollution than the high ability workers and Ω_E is negative. Assuming that Ω_E is negative means that the term *LM* is positive and it has an increasing effect on the shadow price.

The labour market term also involves an effect on redistribution. Negative Ω_E means that when the government seeks to reduce the level of pollution, the wage ratio Ω rises and wage differences decrease. Smaller wage differences mitigate the self-selection constraint and allow the low ability type to earn more. This in turn leads to less need for income transfers and the taxation of a high ability type can be decreased and both types can be better off. The income distribution target in this case is in accordance with environmental targets.

The last term in Eq. (11), G_i, refers to the indirect effect of the externality on government's tax revenues. It tells how much government's tax revenues from commodity taxation change due to the externality. The sign of the public sector term depends on the change in the demand for both goods due to the externality. When an increased pollution level decreases the aggregate demand for both goods[7], the tax revenue effect has a negative sign and the term G_i increases the shadow price. It is also possible that the externality affects the demands for dirty and clean good differently. The determination of the changes in demands would require specifying the exact forms of consumers' utility functions. As in this model utilities are defined only on a very general level, the direction of the demand changes is left undefined.

If a decrease in the level of the externality increases the aggregate demand for goods there will appear a kind of double dividend. From government's point of

---

[6] See e.g. Hamilton and Viscusi (1999), and for a review OECD (2006).

[7] Decreased level of externality means by definition a decreased demand for the dirty good. Thus the increasing aggregate demand implies that the demand for the clean good needs to grow sufficiently due to the effect of the externality. This is possible when dirty good and clean good are substitutes (or at least are not strongly complementary).

view it is advantageous to try to reduce the level of pollution because thus would mean not only better environmental quality but also higher tax revenues from commodity taxes. If the effect of the externality on aggregate demand is negative, emphasising the environmental objective has additional indirect costs through decreased commodity tax revenues.

3.3 The consistency of environmental and redistributional goals

As a whole, the sign of the shadow price is ambiguous. Both of the direct effects increase the shadow price implying the harmfulness of the externality. The interpretation of the indirect effects is in turn less straightforward. On the consumer side the indirect effect of the self-selection constraint is negative when environmental quality and leisure are complements. This means that as the attempt to redistribute increases the incentives to mimic, and as higher level of the externality deters mimicking, in the optimum the governments' marginal harm of the externality is smaller than it would be without redistribution. In this sense, environmental and redistributional objectives are contradictory with respect to this indirect consumer effect. Thus our model implies that there may be a source of conflict between the two goals from the theoretical point of view. However, the labour market effect is more likely to have a positive influence on the shadow price. If this labour market effect is large enough to compensate the negative indirect consumer effect, environmental objectives and redistributional goals are, overall, in accordance.

The assumption that the externality is socially harmful indicates that the shadow price must be positive. It should also be noted that this is not the only conclusion that can be drawn. If the presence of the externality discourages mimicking so much that term C_i is large enough and/or it increases the tax revenues by boosting the demand for goods, it is possible that the shadow price is negative. This implies that the externality would in this case actually be socially beneficial. In the following discussion it will be assumed that the externality is harmful and that its valuation is positive.

4. The optimal taxation

4.1 Marginal tax rates for households

In this model the tax system consists of a direct income tax and an indirect commodity tax. The total tax paid by a worker of type *h* household is the sum of these two taxes $\tau^h(Y) = T(Y^h) + \mathbf{t}^T \mathbf{X}^h$. The marginal effective tax rate, referred to from now on as *METR*, can be found by differentiating total taxes with respect to income, Y^h:

$$\tau^{h\prime}(Y^h) = T' + (1 - T')\mathbf{t}^T \frac{\partial \mathbf{X}^h}{\partial B^h} + \frac{1}{w^h}\mathbf{t}^T \frac{\partial \mathbf{X}^h}{\partial L^h}. \tag{13}$$

The marginal income tax rate is derived from households' optimisation problem. The resulting first order condition is $T' = 1 + \frac{V_L^h}{w^h V_B^h}$. Substituting this into Eq. (13) gives us the *METR* for type *h*

$$\tau^{h\prime} = 1 + \frac{V_L^h}{w^h V_B^h}\left[1 - \mathbf{t}^{\mathbf{T}} \frac{\partial \mathbf{X}^h}{\partial B^h}\right] + \frac{1}{w^h}\mathbf{t}^T \frac{\partial \mathbf{X}^h}{\partial L^h}. \tag{14}$$

The conditional demand functions $\mathbf{X}^h = \begin{bmatrix} X_c^h \\ X_d^h \end{bmatrix}$, *h*=1,2 satisfy the adding-up conditions $\mathbf{q}^T \frac{\partial \mathbf{X}^h}{\partial B^h} = 1$ and $\mathbf{q}^T \frac{\partial \mathbf{X}^h}{\partial Y^h} = 0$. Since $Y^h = w^h L^h$, the second adding-up condition can be written in the form $\mathbf{q}^T \frac{\partial \mathbf{X}^h}{\partial L^h}\frac{1}{w^h} = 0$. Taking these properties into account Eq. (14) becomes

$$\tau^{h\prime} = \frac{1}{w^h}\left[w^h + \frac{V_L^h}{V_B^h}\mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^h} - \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial L^h} \right]. \tag{15}$$

The marginal effective tax rates can be determined from this form by solving term $\frac{V_L^h}{V_B^h}\mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^h}$ from the first order conditions (4)-(7) for both household types respectively.

Consider first the high ability type. The METR for type 2 is

$$\tau'^2 = \frac{1}{w^2}\left[\underbrace{\frac{\mu}{\gamma}\left(\frac{\partial X_d^2}{\partial L^2} - \frac{V_L^2}{V_B^2}\frac{\partial X_d^2}{\partial B^2} \right)}_{\text{term } A_2} + \underbrace{\lambda * \frac{\hat{V}_L^2}{\hat{V}_B^2} L^1 \frac{\partial \Omega}{\partial L^2}}_{\text{term } B_2} \right]. \tag{16}$$

The effective marginal tax rate of type 2 consists of a positive coefficient $\frac{1}{w^2}$ and of two terms A_2 and B_2, where subscripts refer to the type of worker. Thus, the traditional 'no distortion at the top' result does not hold in this framework, resulting from two effects A_2 and B_2.

The term A_2 describes the effect of the externality: if there were no externality, the shadow price would be zero and the whole term would vanish.[8] When the externality is harmful, we notice that term A_2 is positive when the dirty good is a normal good and a substitute for leisure. Intuitively this can be interpreted such that discouraging the labour supply of the high-skilled type becomes desirable, when an increase in income raises the demand for the dirty good thereby worsening the environmental situation.

---

[8] A similar term is also found in Pirttilä and Tuomala (1997).

The term B_2 captures the effect of endogenous wages. Since $\frac{\partial\Omega}{\partial L^2} > 0$, the sign of term B_2 is negative. This implies that when the wage rates are allowed to be adjusted endogenously the effective marginal tax rate of the high ability person decreases: by encouraging the labour supply of type 2, the wage ratio increases, which in turn decreases the incentive to mimic and thus mitigates the otherwise binding self-selection constraint.

As the two terms have opposite signs (when the externality is socially harmful and the dirty good is a normal good and a substitute for leisure), the overall level of the marginal effective tax rate of the high ability type remains ambiguous as long as utility functions and demands are not specified.

Likewise, the effective marginal tax rate for the low ability type is given by

$$\tau'^{1} = \frac{1}{w^1}\left[\underbrace{\frac{\mu}{\gamma}\left(\frac{\partial X_d^1}{\partial L^1} - \frac{V_L^1}{V_B^1}\frac{\partial X_d^1}{\partial B^1}\right)}_{\text{term } A_1} + \underbrace{\lambda^*\left(\frac{\hat{V}_L^2}{\hat{V}_B^2}\Omega - \frac{V_L^1}{V_B^1}\right)}_{\text{term } B_1} + \underbrace{\lambda^*\frac{\hat{V}_L^2}{\hat{V}_B^2}L^1\frac{\partial\Omega}{\partial L^1}}_{\text{term } C_1}\right]. \tag{17}$$

The marginal effective tax rate for the low ability person is determined by a positive coefficient $\frac{1}{w^1}$ and three terms: one referring to the externality, another resulting from the self-selection constraint and the third reflecting the effect of endogenous wages.[9]

The first term, A_1 corresponds to that of the high ability person and is positive if the dirty good is a normal good and a substitute for leisure. The same intuition of the desirability of taxing labour supply as a means to discourage the consumption of the dirty good also holds here. Term B_1 defines the effect of the

---

[9] Note that this distinction is only terminological; the endogenousness of the wage ratio also originates from the self-selection constraint as the changes in relative wages affect the incentives to mimic.

self-selection constraint. The term can be shown to be positive as a result of the single-crossing property[10].

Term C_1 is an outcome of the endogenous wages. It has a positive effect on the *METR*, since $\frac{\partial \Omega}{\partial L^1}$ is negative. This implies that the effect of the wage adjustment increases the marginal effective tax rate for the low ability person. Intuitively, discouraging the labour supply of type 1 decreases the wage ratio which worsens the incentives for mimicking. As all terms, A_1, B_1 and C_1, are positive, the marginal effective tax rate of the low ability type is strictly positive when the dirty good is normal and a substitute for leisure.

The conclusions from Eq. 16 and 17 are that (under the assumptions made) the effective marginal tax rate for the low ability type is strictly positive and the sign of the effective marginal tax rate for the high ability person is ambiguous. The effect of endogenous wages decreases the *METR* of the high productivity worker (term B_2) and raises the *METR* of the low productivity worker (term C_1). If there were no externality, terms A_1 and A_2 would disappear from the equations of marginal effective tax rates. When the dirty good is a normal good and a substitute for leisure and the externality is socially harmful, the presence of the externality increases *METR* for both household types. This result is in accordance with Pirttilä and Tuomala (1997).

An effect similar to terms B_2 and C_1 including the effect of labour supply L^h on the wage ratio has also been reported in other papers on optimal taxes with

---

10 This property states that the indifference curve of the high ability person is flatter than that of the low ability person at a given point. Usually this is presented in income-consumption space in the form $\hat{V}_Y^2 / \hat{V}_B^2 - V_Y^1 / V_B^1 > 0$. Taking a partial derivative of the utility functions with respect to income Y and substituting these derivatives into the usual form of single-crossing condition gives us $\left(\hat{V}_L^2 / \hat{V}_B^2\right)\Omega - V_L^1 / V_B^1 > 0$.

endogenous wages. Aronsson and Sjögren (2003) consider optimal taxation in a case where wages are determined by labour unions. The corresponding effect is captured in their elasticity type term ε_l, but resulting from the imperfect labour market, the sign of the term is not obvious in their framework. Micheletto (2004) also considers marginal effective tax rates with endogenous factor prices and finds that it is desirable to distort the labour supply decisions when, by so doing, government can increase the wage ratio and thus decrease the desirability of mimicking. However, as in his model the wage ratio is a function of factor prices, and there are differences in the intensity of using each labour type, the overall effect of the wage adjustment on the METRs depends on the degree of the complementarity between private good and labour.

4.2 Commodity tax rates

A widely cited result defining the structure of the optimal commodity tax levied on a good creating a harmful externality is Sandmo's (1975) 'additivity property', which was later shown to be a special case of the 'principle of targeting' introduced by Dixit (1985). The principle states that the commodity tax meant to internalise the effect of a harmful externality should be imposed only on the good creating the externality. According to Sandmo's property the internalising part of taxation appears additively in the tax rate. These properties are referred to hereafter as the Sandmo-Dixit principle.

It has been shown that with exogenous wages the Sandmo-Dixit principle continues to hold (Pirttilä and Tuomala, 1997). The issue of the generality of this principle with an assumption of externality is also considered in Kopczuk (2003), who also raises doubts that the principle may not hold in a general equilibrium framework.

The optimal commodity taxation rule can be derived from the first order condition with respect to price level **q**. From Roy's identity we get $V_{\mathbf{q}}^h = -\mathbf{X}^{h^T} V_B^h$

and differentiated consumer's budget constraint implies that $\left[\frac{\partial \mathbf{X}^h}{\partial \mathbf{q}}\right]^T \mathbf{q} = -\mathbf{X}^h$.

Using these, the first order condition (8) can be rewritten in the form[11]

$$\sum_h \mathbf{t}^T \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} = \boldsymbol{\varphi} + \frac{\mu}{\gamma}\sum_h \frac{\partial X_d^h}{\partial \mathbf{q}},$$
$$\text{where } \boldsymbol{\varphi} = -\sum_h \mathbf{X}^{h\,T} + \frac{1}{\gamma}\left[V_B^1 \mathbf{X}^{1^T} + (\delta+\lambda)V_B^2 \mathbf{X}^{2^T} - \lambda \hat{V}_B^2 \hat{\mathbf{X}}^{2^T}\right]. \tag{18}$$

Vector $\boldsymbol{\varphi}$ includes the elements which do not refer to the externality. This part of the commodity taxation thus remains unchanged even if the externality disappears. Commodity taxes can be solved separately by utilizing Cramer's rule:

$$t_c = \frac{1}{J}\left[\sum_h \left(\varphi_c \frac{\partial X_d^h}{\partial q_d} - \varphi_d \frac{\partial X_d^h}{\partial q_c}\right)\right] \text{ and}$$
$$t_d = \frac{1}{J}\left[\sum_h \left(\varphi_d \frac{\partial X_c^h}{\partial q_c} - \varphi_c \frac{\partial X_c^h}{\partial q_d}\right)\right] + \frac{\mu}{\gamma}, \tag{19}$$

where J is the determinant of matrix $\left[\sum_h \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}}\right]^T$. The part of the commodity tax indicating the harmfulness of the externality is the shadow price $\frac{\mu}{\gamma}$. Dixit's (1985) general result, the principle of targeting, continues to hold under the assumption of endogenous wages, as the externality-based part of the commodity tax only appears in the tax rate of the dirty good. A more specific result is Sandmo's additivity property, which states that the term connected to the harmful externality appears additively in the tax rate of the dirty good. This property is also satisfied, so the Sandmo-Dixit principle can be generalised to the case of endogenous wages.

[11] A more detailed derivation is presented in Appendix B.

The optimal level of commodity taxes can thus be written in the same form as in the case of exogenous wages (Pirttilä and Tuomala, 1997). However, the tax rate of the dirty good is not the same, as the shadow price of the externality $\frac{\mu}{\gamma}$ differs from the shadow price in the case of exogenous wages by the labour market term (LM) and direct production side effect (P_d). Thus, the tax rates have changed due to the endogeneity assumption, but the general form of the optimal tax rates has remained the same.

In technical terms, it is not surprising that Sandmo-Dixit principle continues to hold, as the wage ratio Ω does not depend on the consumer prices. However, if the wage ratio could be affected by the commodity tax rates, this result might change.

5. Conclusions

This paper analysed optimal tax policy in the presence of a harmful externality under an assumption of endogenous wages. In a two-type economy with high and low productivity households government has to take the self-selection constraint into account when designing the optimal tax policy. As government pursues its environmental and redistributive objectives, it is assumed to have non-linear income taxation and linear commodity taxation as instruments.

An essential term in the optimal taxes is the social valuation of the externality. This valuation consists of direct and indirect effects on consumers depending on the marginal willingnesses to pay to avoid the externality, direct effect on producers resulting from the influence of the externality on production, the labour market effect affecting indirectly both producers and consumers through the adjustment of the wage ratio, the change in the tax revenues affecting government and the environmental feedback parameter. Both direct effects increase the valuation of the externality reflecting its harmfulness on both consumers and producers. The self-selection effect is negative on the social valuation of the externality when environmental quality and leisure are complementary. This means that environmental deterioration may discourage mimicking and promote government's redistribution goals, i.e. environmental and redistributional objectives may be contradictory. However, on the producer side the direct effect and the indirect effect through wage adjustment on the labour market both have positive signs (under reasonable assumptions), implying that contributing to environmental quality also promotes redistribution. Thus, the overall effect is ambiguous.

Introducing the externality results in an additional term in the marginal effective tax rates of both types. This term is positive, i.e. it tends to increase the effective marginal tax rates of both types when the externality is harmful (indicated by a positive shadow price) and when the dirty good is a normal good and a substitute for leisure. The extra term reflecting the effect of the endogenous wages decreases the effective marginal tax rate of the high productivity worker and raises that of the low productivity worker. In optimal commodity taxation the Sandmo-Dixit principle continues to hold: the externality internalising part, which is equal to the social valuation of the externality, appears only in the tax rate of the dirty good and enters the tax rate additively.

References

Aronsson, T. and T. Sjögren (2003), 'Income Taxation, Commodity Taxation, and Provision of Public Goods under Labor Market Distortions', *FinanzArchiv* 59, 347-370.

Atkinson, A. B. and J. E. Stiglitz (1976), 'The Design of Tax Structure: Direct versus Indirect Taxation', *Journal of Public Economics* 10, 55-75.

Brännlund, R. and J. Nordström (2004), 'Carbon Tax Simulations Using a Household Demand Model', *European Economic Review* 48, 311-333.

Diamond, P.A. and J. Mirrlees (1971), 'Optimal Taxation and Public Production I: Production Efficiency and II: Tax Rules', *American Economic Review* 61: 8-27 and 261-278.

Dixit, A. (1985), 'A Tax Policy in Open Economies', in: A.J. Auerbach and M. Feldstein (eds.), '*Handbook of Public Economics, vol. I*', North-Holland, Amsterdam. 313-374

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

Hamilton, J. T. and W. K. Viscusi (1999), 'Calculating Risks? The Spatial and Political Dimensionas of Hazardous Waste Policy', MIT Press, Cambridge.

Harrison, D. Jr. (1994), 'The Distributive Effects of Economic Instruments for Environmental Policy', Paris: OECD.

Kopczuk, W. (2003), 'A Note on Optimal Taxation in the Presence of Externalities', *Economics Letters* 80, 81-86.

Micheletto, L. (2004), 'Optimal Redistributive Policy with Endogenous Wages', *FinanzArchiv* 60, 141-159.

Mirrlees, J. A. (1976), 'Optimal Tax Theory: A Synthesis', *Journal of Public Economics* 20, 77-87.

Naito, H. (1999), 'Re-examination of Uniform Commodity Taxes under a Non-linear Income Tax System and Its Implication for Production Efficiency', *Journal of Public Economics* 71, 165-188.

OECD (2006), '*The Distributional Effects of Environmental Policy*', Y. Serret and N. Johnstone (eds.). Edwar Elgar Publishing.

Pirttilä, J. and M. Tuomala (1997), 'Income Tax, Commodity Tax and Environmental Policy', *International Tax and Public Finance* 4, 379-393.

Samuelson, P. (1954), 'The Pure Theory of Public Expenditure', *Review of Economics and Statistics*, 36, 387-389.

Sandmo, A. (1975), 'Optimal Taxation in the Presence of Externalities', *Swedish Journal of Economics* 77, 86-98.

Sandmo, A. (1980), 'Anomaly and Stability in the Theory of Externalities', *Quarterly Journal of Economics* 94, 799-807.

Smith, S. (1992), 'Taxation and the Environment: A Survey', *Fiscal Studies* 13, 21-57.

Stern, N. H. (1982), 'Optimum 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.

West, S. and R. C. Williams III (2004), ‘Estimates from a Consumer Demand System: Implications for the Incidence of Environmental Taxes’, *Journal of Environmental Economics and Management* 47, 535-558.

APPENDIX A: Derivation of the shadow price of the externality

First add and subtract $\lambda \hat{V}_B^2 \dfrac{V_E^1}{V_B^1}$ from Eq. (9) to obtain

$$\begin{aligned}&\left(V_B^1 - \lambda \hat{V}_B^2\right)\frac{V_E^1}{V_B^1} + (\delta + \lambda)V_B^2 \frac{V_E^2}{V_B^2} - \lambda \hat{V}_B^2 \left(\frac{\hat{V}_E^2}{\hat{V}_B^2} - \frac{V_E^1}{V_B^1}\right) \\ &\qquad - \lambda \hat{V}_L^2 L^1 \frac{\partial \Omega}{\partial E} + \gamma\left(F_E - \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E}\right) + \mu\left(1 - \sum_h \frac{\partial X_d^h}{\partial E}\right) = 0.\end{aligned} \tag{23}$$

Let $MWP_{EB}^h = -\dfrac{V_E^h}{V_B^h}$ denote the marginal willingness to pay to avoid the externality. Substituting this into the previous equation yields

$$\begin{aligned}&-\left(V_B^1 - \lambda \hat{V}_B^2\right)MWP_{EB}^1 - (\delta + \lambda)V_B^2 MWP_{EB}^2 + \lambda \hat{V}_B^2\left(M\hat{W}P_{EB}^2 - MWP_{EB}^1\right) \\ &\qquad - \lambda \hat{V}_L^2 L^1 \frac{\partial \Omega}{\partial E} + \gamma\left(F_E - \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E}\right) + \mu\left(1 - \sum_h \frac{\partial X_d^h}{\partial E}\right) = 0.\end{aligned} \tag{24}$$

Substituting into the first order conditions (5) and (7) we obtain

$$\begin{aligned}&-\left(\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial B^1} + \mu \frac{\partial X_d^1}{\partial B^1}\right)MWP_{EB}^1 - \left(\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2} + \mu \frac{\partial X_d^2}{\partial B^2}\right)MWP_{EB}^2 \\ &\qquad + \lambda \hat{V}_B^2\left(M\hat{W}P_{EB}^2 - MWP_{EB}^1\right) - \lambda \hat{V}_L^2 L^1 \frac{\partial \Omega}{\partial E} \\ &\qquad\qquad + \gamma\left(F_E - \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E}\right) + \mu\left(1 - \sum_h \frac{\partial X_d^h}{\partial E}\right) = 0.\end{aligned} \tag{25}$$

Next we use the Slutsky-type property of the conditional demand function to simplify the equation further. The conditional demand function x_i^h is obtained by minimizing the consumer's cost and keeping the utility constant. For the conditional demand function it holds that $\frac{\partial X_i^h}{\partial E} = \frac{\partial x_i^h}{\partial E} - MWP_{EB}^h \frac{\partial X_i^h}{\partial B^h}$. The term $\sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E}$ can now be written as $\sum_h \mathbf{p}^T \frac{\partial \mathbf{x}^h}{\partial E} - \sum_h MWP_{EB}^h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^h}$. Edwards, Keen and Tuomala (1994) also show that $\mathbf{p}^T \frac{\partial \mathbf{x}^h}{\partial E} = MWP_{EB}^h - \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E}$. Substituting these into the terms in the last row of Eq. (25) and using a definition $\lambda^* = \frac{\lambda \hat{V}_B^2}{\gamma}$ we obtain

$$
\begin{aligned}
&-\frac{\mu}{\gamma}\sum_h MWP_{EB}^h \frac{\partial X_d^h}{\partial B^h} + \lambda^*\left(M\hat{W}P_{EB}^2 - MWP_{EB}^1\right) + F_E \\
&\quad + \sum_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E} - \frac{\lambda}{\gamma}\hat{V}_L^2 L^1 \frac{\partial \Omega}{\partial E} - \sum_h MWP_{EB}^h \\
&\quad\quad + \frac{\mu}{\gamma}\left(1 - \sum_h\left(\frac{\partial x_d^h}{\partial E} - MWP_{EB}^h \frac{\partial X_d^h}{\partial B^h}\right)\right) = 0.
\end{aligned}
\tag{26}
$$

In the equation above there are several terms that can be cancelled out and the equation can be simplified further to the form

$$
\begin{aligned}
\frac{\mu}{\gamma}\left(1 - \sum_h \frac{\partial x_d^h}{\partial E}\right) &= \sum_h MWP_{EB}^h - \lambda^*\left(M\hat{W}P_{EB}^2 - MWP_{EB}^1\right) \\
&\quad - F_E + \lambda^* \frac{\hat{V}_L^2}{\hat{V}_B^2} L^1 \frac{\partial \Omega}{\partial E} - \sum_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E}.
\end{aligned}
\tag{27}
$$

From this form the shadow price of the externality can be solved, and it can be written as in Eq. (11).

APPENDIX B: Derivation of commodity tax rates

First rewrite the first order condition (8) with the help of Roy's identity $V_{\mathbf{q}}^h = -\mathbf{X}^{h^T} V_B^h$ as

$$\begin{aligned} & -V_B^1 \mathbf{X}^{1^T} - (\delta + \lambda) V_B^2 \mathbf{X}^{2^T} + \lambda \hat{V}_B^2 \hat{\mathbf{X}}^{2^T} \\ & \qquad - \gamma \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} - \mu \sum_h \frac{\partial X_d^h}{\partial \mathbf{q}} = \mathbf{0} \end{aligned} \tag{28}$$

By differentiating consumer's budget constraint $\mathbf{q}^T \mathbf{X}^h = B^h$ with respect to the price vector $\mathbf{q}$ we obtain a condition $\left[\frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} \right]^T \mathbf{q} + \mathbf{X}^h = \mathbf{0}^T$, which is in transposed into the form $\mathbf{q}^T \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} = -\mathbf{X}^{h^T}$. Substituting this condition into Eq. (28) and rearranging some terms it can be written as

$$\begin{aligned} & -\boldsymbol{\varphi} + \sum_h \mathbf{t}^T \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} - \frac{\mu}{\gamma} \sum_h \frac{\partial X_d^h}{\partial \mathbf{q}} = \mathbf{0}, \\ & \text{where } \boldsymbol{\varphi} = -\sum_h \mathbf{X}^{h^T} + \frac{1}{\gamma} \left[V_B^1 \mathbf{X}^{1^T} + (\delta + \lambda) V_B^2 \mathbf{X}^{2^T} - \lambda \hat{V}_B^2 \hat{\mathbf{X}}^{2^T} \right] \end{aligned} \tag{29}$$

The vector $\boldsymbol{\varphi}$ includes the elements not referring to the externality. This part of the commodity taxation thus remains unchanged even if the externality disappears.

To solve the optimal commodity tax rate, we write Eq. (29) in the form $\sum_h \mathbf{t}^T \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} = \boldsymbol{\varphi} + \frac{\mu}{\gamma} \sum_h \frac{\partial X_d^h}{\partial \mathbf{q}}$, which is equivalent to a transposed matrix form

$$\begin{bmatrix} \sum_h \frac{\partial X_c^h}{\partial q_c} & \sum_h \frac{\partial X_d^h}{\partial q_c} \\ \sum_h \frac{\partial X_c^h}{\partial q_d} & \sum_h \frac{\partial X_d^h}{\partial q_d} \end{bmatrix} \begin{bmatrix} t_c \\ t_d \end{bmatrix} = \begin{bmatrix} \varphi_c \\ \varphi_d \end{bmatrix} + \frac{\mu}{\gamma} \begin{bmatrix} \sum_h \frac{\partial X_d^h}{\partial q_c} \\ \sum_h \frac{\partial X_d^h}{\partial q_d} \end{bmatrix}, \tag{30}$$

from which t_c and t_d can be solved by Cramer's rule.

Chapter 3
Optimal Tax Policy with Environmental Externalities and Heterogeneous Preferences*

Sanna Tenhunen†
University of Tampere and FDPE

Abstract

This study considers simultaneously two important aspects of taxation: environmental policy and redistribution. Tax policy is constrained by the asymmetric information on agents' productivities and preferences. Two-dimensional heterogeneity affects the optimality of commodity taxation: it can be used to redistributional for environmental purposes, but there seems to be a trade-off between these objectives. However, the contradiction between the two aspects is not as clear as in the case with identical consumer preferences.

It is also shown that the Sandmo-Dixit result of the separability of environmental taxes fails with two-dimensional heterogeneity in the pooling optimum, but not in the separating optimum. The explanation for this is that there are too few policy instruments in the pooling equilibrium: commodity taxes should take care of both redistribution and externality internalisation.

Keywords: heterogeneous preferences, externalities, commodity taxation

JEL: H21, H23

*I am grateful to Thomas Aronsson, Vidar Christiansen, Jukka Pirttilä and Matti Tuomala for the most helpful comments.

†E-mail: sanna.tenhunen@uta.fi

1 Introduction

In the recent literature on optimal taxation and redistribution the assumption of preference homogeneity is abandoned and households are allowed to have different tastes. Heterogeneity can arise from different tastes for leisure, different tastes for consumption or from some combination of these. Most of the literature concentrates on the former type of heterogeneity[1]. Introducing an additional dimension to individuals' characteristics makes the analysis more difficult: there are now four groups of households. To make the calculations more reasonable the number of household types is often reduced to three. Three groups are adequate in most cases to characterise the essential effects of heterogeneous preferences.

To design the optimal tax scheme, the direction of redistribution needs to be determined. It is no longer straightforward to determine which group should be treated the most gently. Sandmo (1993) discusses the problem of comparing utilities when tastes differ. He finds that even if utilities could be compared, a change in the parameter set describing the economy (such as prices) can reverse the order. The literature on social choice also devotes attention to this question[2]. The social welfare approach aims at taking individuals' different preferences into account even when it means treating individuals with equal skill levels differently. According to the horizontal equity principle, individuals should not be treated differently on grounds of their different preferences. However, under asymmetric information this may be a costly or even an impossible requirement.

In a model with heterogeneous and usually unobservable preferences the possibility of screening is of special interest. Heterogeneous preferences were first analysed in the framework considering the optimality of workfare e.g. by Besley and Coate (1995), Beaudry and Blackorby (1997) and Cuff (2000). Tarkiainen and Tuomala (1999) study optimal tax policy numerically in an economy where households differ with respect to abilities and work preferences. Boadway et al. (2002) find that if there are some important but unobservable

[1] An exception is Blomquist and Christiansen (2004), where heterogeneity in preferences for consumption goods is also discussed.

[2] See Fleurbaey and Maniquet (1999) for a survey of this subject.

characteristics affecting individual's choices, it can be optimal to abandon the income tax schedule based on equal weights on preference groups, even if the resulting tax scheme seems regressive in terms of the observable characteristics, such as income.

With two-dimensional unobservable characteristics any action revealing characteristics under asymmetric information is valuable. Introducing an additional consumption good to the economy offers an opportunity to study the role of indirect taxation. Saez (2002) finds that a tax on commodity is desirable when individuals with high income have a relatively higher preference for the commodity or if the consumption of the commodity increases with leisure. Jordahl and Micheletto (2005) also report parallel results: in a model with heterogeneous preferences the consumption of a good complementary to leisure does not need to be discouraged by taxation, and a commodity that is expected to be encouraged should not always be subsidised by a negative tax rate. Blomquist and Christiansen (2004) suggest that commodity taxation may get a new role as a device for differentiating between different groups. Contrary to the Atkinson-Stiglitz result on the redundancy of commodity taxation, when individuals have different preferences for consumption goods imposing a tax on commodities may be desirable even when preferences are separable.

In the light of the earlier studies an assumption of heterogeneous preferences seems to affect the recommendations of optimal tax policy. This study aims at combining two fields of research: the recent extension of the optimal tax models to heterogeneous preferences and a somewhat older discussion of the redistributional problems of environmental taxation[3]. In the earlier literature on optimal taxation under environmental externalities agents are usually assumed to have homogeneous preferences. A model with mixed taxation with one-dimensional heterogeneity and environmental externalities is used e.g. in Pirttilä and Tuomala (1997). Aronsson (2005) studies environmental policy and taxation with an emphasis on the employment aspect. An exception is

[3] An excellent survey of environmental policy as a part of optimal taxation can be found in Bovenberg and Goulder (2002). For a survey of problems in combining environmental and redistributional aspects see e.g. Smith (1992), Harrison (1994) and OECD (2001). More recent research on the possible regressivity of environmental taxes can be found in Walls and Hanson (1999) and Jacobsen, Birr-Pedersen and Wier (2003).

Cremer, Gahvari and Ladoux (1998, 2001), who consider optimal taxation in a case of two-dimensional heterogeneity and environmental externalities, but they leave the redistributional aspect in the background. The assumption that agents have different preferences for leisure seems very realistic. There is also very likely to be a dependence in the preference for leisure and environmental quality. For example, people enjoying spending their time in a summer cottage in the countryside may put more emphasis on environmental quality than someone prefering less leisure.[4] Thus, environmentally related taxes probably have different effects on different consumers' behaviour. There is also a chance that heterogeneous preferences may serve as a screening tool in tax policy. Our aim here is to take a step towards a more general setting by introducing heterogeneous preferences to optimal taxation under environmental externalities.

A closer analysis of a term defining the valuation of environmental externality reveals that there are some terms induced by the redistribution constraint that have the opposite effect compared to environmental aims. Although the existence of such terms does not imply that environmental and redistributional aspects are contradictory, it can be interpreted as one possible explanation for the empirical findings of the regressivity of environmental taxes.

The model used here assumes mixed taxation, i.e. a non-linear income tax and a linear commodity tax. The valuation for the environmental externality and the optimal commodity tax rule will be considered in two cases: in the separating equilibrium, where each household choose their own income-consumption bundle and in the pooling optimum, where two of the households cannot be distinguished by thei choices. Thus, unlike Cremer, Gahvari and Ladoux (1998, 2001), we consider optimal tax policy also with redistributional aspect and the possibility of different types of optima.

The paper is constructed as follows. In Section 2 we introduce the model. Section 3 derives the optimisation problem and the first order conditions. The valuation of the externality is derived in Section 4, where the trade-off between environmental and redistributional aspects is also discussed. Section 5

[4] Although this case of complementarity seems more plausible to us, we realise that it is naturally only one option: environmental quality and leisure may also be substitutes, or the relationship might even be non-monotonic.

concentrates on the commodity tax rules and Sandmo-Dixit principle. Finally, Section 6 concludes.

2 The model

The model here is very similar to that used in Blomquist and Christiansen (2004) with an extension to a harmful environmental externality. It is based on the Mirrlees (1976) type of optimal income tax model with heterogeneous households, differing with respect to productivities (as in Stiglitz, 1982 and Stern, 1982) and preferences.

To avoid excessively restrictive assumptions of utilities, we assume that heterogeneity results from the different preferences for leisure. However, unless preferences are completely separable, the amount of leisure is also likely to affect households' consumption and their assessment of environmental quality. The relation depends on the complementarity of leisure and consumption and environmental quality. When leisure and environmental quality are complements, the household with a stronger preference for leisure also has a stronger preference for the environment. Thus heterogeneity is also reflected to preferences for consumption and environment.

Here we assume three types of households as characterised in Table 1 below.

		productivity	
		low	*high*
preference	*strong*	-	type 2
for leisure	*weak*	type 1	type 3

Table 1: Characteristics of household groups

Households supply labour L^h and receive an exogenous wage rate w^h, which reflects their productivities, i.e. $w^1 < w^2 = w^3$. There is a constant wage ratio $\Omega = \frac{w^1}{w^3}$. Labour income $Y^h = w^h L^h$ is taxed by the optimal non-linear income tax scheme $T(Y^h)$. We also assume that the labour income of type 1 households is the lowest and the gross income of type 3 households is the highest. Households use all their net income $B^h = Y^h - T(Y^h)$ for consumption. There are two goods in the markets[5] denoted by a matrix $\mathbf{X} = \begin{bmatrix} X_c \\ X_d \end{bmatrix}$, where X_c is the "clean" good, whereas X_d is the "dirty" good creating a harmful environmental externality. Both goods are assumed to be normal. Households' indirect utility function is given by $V(\mathbf{q}, B, L, E)$, which is obtained by maximising the direct utility $U(\mathbf{X}, L, E)$ subject to the budget constraint $\mathbf{q}^T\mathbf{X} = B$, with T denoting a matrix transpose.

There is a linear commodity tax $\mathbf{t} = \begin{bmatrix} t_c \\ t_d \end{bmatrix}$, so consumer prices can be denoted by a vector $\mathbf{q} = \mathbf{t} + \mathbf{p}$, where $\mathbf{p} = \begin{bmatrix} p_c \\ p_d \end{bmatrix}$ stands for producer prices. The demand $X_i^h\left(\mathbf{q}, B^h, L^h, E\right)$ of household h $(h = 1, 2, 3)$ for commodity i $(i = c, d)$ is a function of prices $\mathbf{q}$, after-tax income B^h, labour supply L^h and externality E, where $E = \sum_h X_d^h\left(\mathbf{q}, B^h, L^h, E\right)$.

The public sector has two preferences: more equal income distribution[6] and cleaner environment. It has two tax devices, non-linear income tax and linear commodity tax, to finance a constant revenue requirement $\bar{G}$. The redistribution from the high productivity households to the low productivity households is constrained by self-selection constraints as productivities are unobservable. Furthermore, there is a problem in distinguishing between the households receiving low income due to low productivity and the high productivity households earning less due to their strong preference for leisure. If income is redistributed by means of income taxation from type 3 households to both type 1 and type 2 households, the horizontal equity principle demanding an equal treatment for households with the same characteristics is violated. Thus income taxation has to be designed so that all households choose the combination of labour supply and net income meant for them rather than mimicking the choice of other household type.

In the separating optimum all households choose a different point. We

[5] We omit the production side of the economy because it does not have any effect on our results.

[6] We implicitly assume that the current distribution of income is sufficiently unequal for redistribution to be desirable. However, the results derived here would also be applicable to the opposite case.

assume that the binding self-selection constraints are, firstly, a type 3 household should not mimic a type 2 household, and secondly, a type 2 household should not mimic a type 1 household. In mathematical form the constraints are given as $V^3(\mathbf{q}, B^3, L^3, E) \geq V^3(\mathbf{q}, B^2, L^2, E)$ and $V^2(\mathbf{q}, B^2, L^2, E) \geq V^2(\mathbf{q}, B^1, \Omega L^1, E)$.

Not all self-selection constraints are necessarily binding. In the pooling optimum we assume that households of types 1 and 2 voluntarily choose the same point and thus income taxation cannot be used to differentiate between these two groups[7]. However, although the observable net consumption of the pooled households is equal, they do not necessarily choose exactly the same consumption bundles, as they are assumed to have heterogeneous preferences. This raises a question of whether commodity taxes could be used as a screening tool for redistributional purposes. There are again two self-selection constraints $V^3(\mathbf{q}, B^3, L^3, E) \geq V^3(\mathbf{q}, B^2, L^2, E)$ and $V^3(\mathbf{q}, B^3, L^3, E) \geq V^3(\mathbf{q}, B^1, \Omega L^1, E)$ but in the pooling case, as we have $B^1 = B^2$, $Y^1 = Y^2$ and $\Omega L^1 = L^2$, they represent exactly the same outcome. Naturally, when households 1 and 2 voluntarily choose the same income consumption bundle, they do not have any incentive to mimic each other.

3 The optimisation problem

Government optimises the utility of the low productivity (type 1) household subject to two Pareto constraints

$$V^2\left(\mathbf{q}, B^1, L^1, E\right) \geqslant \bar{V}^2 \text{ and } V^3\left(\mathbf{q}, B^3, L^3, E\right) \geqslant \bar{V}^3 \tag{1}$$

[7] We chose this pooling equilibrium because it seemed most plausible. If types 2 and 3 pooled, they would have identical amount of leisure, as they have the same productivities. Because type 2 prefers more leisure, he will probably not be willing to choose the same bundle as type 3. The other possible pooling optimum, where types 1 and 3 are pooled also seems unlikely. Types 1 and 3 have the same preference for leisure, but type 3 has higher productivity. If type 3 chooses the same bundle as type 1, then so does type 2, as he enjoys more the extra leisure achieved as a result of higher productivity. Thus, all households would end up at the same point, which would make differentiated tax policy redundant.

In the separating optimum there are two self-selection constraints, given by

$$V^3\left(\mathbf{q}, B^3, L^3, E\right) \geq V^3\left(\mathbf{q}, B^2, L^2, E\right) \text{ and } V^2\left(\mathbf{q}, B^2, L^2, E\right) \geq V^2\left(\mathbf{q}, B^1, \Omega L^1, E\right) \quad (2)$$

whereas in the pooling case one constraint $V^3(\mathbf{q}, B^3, L^3, E) - V^3(\mathbf{q}, B^1, \Omega L^1, E)$ is sufficient to capture the restriction of mimicking. Another restriction comes from government's budget constraint requiring that the income from taxes equal the revenue requirement. Using the consumer's budget constraint this can be rewritten as

$$\sum_h Y^h - \sum_h \mathbf{p}^T \mathbf{X}^h\left(\mathbf{q}, B^h, L^h, E\right) = \bar{G} \quad (3)$$

And finally the fifth constraint captures the externality defined as the aggregate consumption of the dirty good

$$\sum_h X_d^h\left(\mathbf{q}, B^h, L^h, E\right) = E \quad (4)$$

Because the Lagrange function and the first order conditions are very similar in two optima, we present only the pooling case here, whereas the optimisation conditions in the separating equilibrium appear in Appendix A. The Lagrangean of the optimisation problem in the pooling case is given by

$$\begin{aligned}\Psi = {} & V^1\left(\mathbf{q}, B^1, L^1, E\right) \\ & + \delta_2\left[V^2\left(\mathbf{q}, B^1, \Omega L^1, E\right) - \bar{V}^2\right] + \delta_3\left[V^3\left(\mathbf{q}, B^3, L^3, E\right) - \bar{V}^3\right] \\ & + \lambda\left[V^3\left(\mathbf{q}, B^3, L^3, E\right) - V^3\left(\mathbf{q}, B^1, \Omega L^1, E\right)\right] \\ & + \gamma\left[\sum_h w^h L^h - \sum_h \mathbf{p}^T \mathbf{X}^h\left(\mathbf{q}, B^h, L^h, E\right) - \bar{G}\right] \\ & + \mu\left[E - \sum_h X_d^h\left(\mathbf{q}, B^h, L^h, E\right)\right] \quad (5)\end{aligned}$$

The optimisation problem is to choose B^1, B^3, L^1, L^3, E and $\mathbf{t}$ optimally. The first order conditions needed are

$$V_L^1 + \delta_2 V_L^2 \Omega - \lambda \widehat{V}_L^3 \Omega + \gamma \left(\sum_{h=1,2} w^h - \sum_{h=1,2} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial L^1} \right) - \mu \sum_{h=1,2} \frac{\partial X_d^h}{\partial L^1} = 0 \quad (6)$$

$$V_B^1 + \delta_2 V_B^2 - \lambda \widehat{V}_B^3 - \gamma \sum_{h=1,2} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^1} - \mu \sum_{h=1,2} \frac{\partial X_d^h}{\partial B^1} = 0 \quad (7)$$

$$(\delta_3 + \lambda) V_L^3 + \gamma \left(w^3 - \mathbf{p}^T \frac{\partial \mathbf{X}^3}{\partial L^3} \right) - \mu \frac{\partial X_d^3}{\partial L^3} = 0 \quad (8)$$

$$(\delta_3 + \lambda) V_B^3 - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^3}{\partial B^3} - \mu \frac{\partial X_d^3}{\partial B^3} = 0 \quad (9)$$

$$V_E^1 + \delta_2 V_E^2 + (\delta_3 + \lambda) V_E^3 - \lambda \widehat{V}_E^3 - \gamma \sum_{h=1,2,3} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E} + \mu \left(1 - \sum_{h=1,2,3} \frac{\partial X_d^h}{\partial E} \right) = 0 \quad (10)$$

$$V_{\mathbf{q}}^1 + \delta_2 V_{\mathbf{q}}^2 + (\delta_3 + \lambda) V_{\mathbf{q}}^3 - \lambda \widehat{V}_{\mathbf{q}}^3 - \gamma \sum_{h=1,2,3} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} - \mu \sum_{h=1,2,3} \frac{\partial X_d^h}{\partial \mathbf{q}} = 0 \quad (11)$$

where the hat terms refer to mimickers, i.e. true type 3 household mimicking type 1 household.

4 The harmfulness of the externality

The valuation of the externality is a useful term for two reasons. First, the optimal tax rules depend on the environmental quality via the term defining the valuation of the externality. Second, a closer look at the terms in it enables a simultaneous consideration of environmental and redistributional aspects. The valuation of the environmental externality shows how much harm externality produces and it is measured here by the shadow price. Here the form $\frac{\mu}{\gamma}$ is used, i.e. the shadow price is given relative to the government's tax revenues.

We note that by taking a derivative of the Lagrangian with respect to B^2, we obtain a useful term $\left[\frac{\delta_2 V_B^2}{\gamma} - \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2} - \frac{\mu}{\gamma}\frac{\partial X_d^2}{\partial B^2}\right] = \frac{1}{\gamma}\frac{\partial \Psi}{\partial B^2} = \frac{\Psi_{B^2}}{\gamma}$. This term indicates the value given to the hypothetical increase in the net income of type 2 households. A similar term was utilised in Blomquist and Christiansen (2004). It can be interpreted to be negative if the desired direction of redistribution is from the high ability households towards the low ability households.

The implicit form[8] of the shadow price can be solved from Equation (10) (for derivation, see Appendix B)

$$\frac{\mu}{\gamma} = \sigma \left[\begin{array}{c} \sum\limits_h MWP_{EB}^h + \frac{\Psi_{B^2}}{\gamma}\left[MWP_{EB}^2 - MWP_{EB}^1\right] \\ -\lambda^* \left[\widehat{MWP}_{EB}^3 - MWP_{EB}^1\right] - \sum\limits_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E} \end{array} \right] \tag{12}$$

where $-\frac{V_E^h}{V_B^h} = MWP_{EB}^h$ is the marginal willingness to pay to avoid the externality, $\lambda^* = \frac{\lambda \widehat{V}_B^3}{\gamma}$[9] and $\mathbf{x}^h$ is the compensated conditional demand function

[8] Written in this form Ψ_{B^2} contains the shadow price $\frac{\mu}{\gamma}$. In explicit form the harmfulness of the externality is given by

$$\frac{\mu}{\gamma} = \sigma^* \left\{ \begin{array}{c} \sum\limits_h MWP_{EB}^h + \left[\frac{\delta_2 V_B^2}{\gamma} - \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^1}\right]\left[MWP_{EB}^2 - MWP_{EB}^1\right] \\ -\lambda^* \left[M\hat{W}P_{EB}^3 - MWP_{EB}^1\right] - \sum\limits_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E} \end{array} \right\}$$

where $\sigma^* \frac{1}{1 - \sum\limits_h \frac{\partial x_d^h}{\partial E} + \frac{\partial X_d^2}{\partial B^2}\left[MWP_{EB}^2 - MWP_{EB}^1\right]}$ and $\lambda^* = \frac{\lambda \widehat{V}_B^3}{\gamma}$.

Here coefficient σ^* is positive; it only has an additional positive term in the denominator compared to the ordinary environmental feedback parameter σ. The conclusions of the terms are identical, as $\Psi_{B^2} < 0$ implies that coefficient $\frac{\delta_2 V_B^2}{\gamma} - \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^1}$ is also negative.

[9] The coefficient λ^* captures the shadow price of the self-selection constraint taking into account both total welfare through term λ/γ and the private utility of the mimicker through term $\widehat{V}_B^3$.

for type h given by minimising households' expenditure subject to keeping the direct utility $U^h = U(\mathbf{X}^h, L^h, E)$ on a given level u^h: i.e. $\mathbf{x}^h(\mathbf{q}, Y, u, E, w) = \arg\min_{\mathbf{X}^h} \{\mathbf{q}^T\mathbf{X}^h \,|U^h \geqslant u^h\}$. The coefficient $\sigma = \frac{1}{1-\sum_h \frac{\partial x_d^h}{\partial E}}$ can be interpreted as the environmental feedback parameter, following Sandmo (1980).

The exact sign of the shadow price cannot be determined from the form we have. However, as long as the externality is harmful, the shadow price can be assumed to be positive. Separate consideration of the terms allows us to examine the interaction between environmental and redistributional aspects.

The environmental feedback parameter σ is known to be positive (Sandmo, 1980). The first term in brackets is also positive, implying the direct harm of the externality. The last term referring to government's tax revenues from commodity taxes depends on how the externality affects the demand for goods. The first and the last terms are similar to those in the earlier literature (see Pirttilä and Tuomala, 1997; and Tenhunen, 2004), so we concentrate here on the two terms in the middle.

MWP_{EB}^h depends on how much leisure a household has: it increases with leisure, when environmental quality and leisure are complements[10]. From the two terms in the middle with differences in MWP_{EB}^h, now only $\widehat{MWP}_{EB}^3 - MWP_{EB}^1$ refers to mimicking. When mimicking, high productivity households (type 3) have more leisure than true type 1 households with lower productivity. When the environmental harm is decreased by lowering the level of the externality, mimicking becomes more attractive. To prevent this, the income tax of type 3 households has to be lowered and redistribution becomes more unequal. Thus, environmental and redistributional terms lead on opposite directions, implying a trade-off between the two government preferences in this case. A similar result was found earlier e.g. in Pirttilä and Tuomala (1997).

The second term in brackets, the difference in the marginal willingnesses to pay, $MWP_{EB}^2 - MWP_{EB}^1$ is positive, as both households 1 and 2 choose the same income level but the type 2 household has higher productivity and

[10]Many of the conclusions in this paper are based on the assumption of environmental quality and leisure being complements. The analysis in the case of substitutes would go through analogously. As the former case seems more plausible to us, to avoid confusion we concentrate on the case where the harmfulness of the externality increases with leisure.

thus more leisure. This term does not refer to mimicking nor to redistribution, because in the pooling optimum these two household types cannot be differentiated and thus no redistribution can be made by income taxation. The term rather refers to the difference in the preferences for environmental quality. The fact that households are not equal affects the harmfulness of the shadow price. With negative Ψ_{B^2} the difference between MWP^2_{EB} and MWP^1_{EB} alleviates the harmful effect of the externality. Thus this term affects in the same direction as the third term referring to redistribution. If an increase in the level of the externality widens the difference between the marginal willingnesses to pay, some of the harm from the externality is compensated by gains from redistribution. This trade-off between environmental quality and redistribution indicates that a higher level of the environmental externality can be used as a tool to deter mimicking and to redistribute income between otherwise indistinguishable households.

In the separating optimum the shadow price is given by[11]

$$\frac{\mu}{\gamma} = \sigma \left[\begin{array}{c} \sum\limits_h MWP^h_{EB} - \lambda^*_2 \left[\widehat{MWP}^2_{EB} - MWP^1_{EB} \right] \\ -\lambda^*_3 \left[\widehat{MWP}^3_{EB} - MWP^2_{EB} \right] - \sum\limits_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E} \end{array} \right], \tag{13}$$

where $\lambda^*_h = \frac{\lambda \widehat{V}^h_B}{\gamma}, h = 2, 3$, and, as earlier, $MWP^h_{EB} = -\frac{V^h_E}{V^h_B}$ is the marginal willingness to pay to avoid the externality and $\sigma = \frac{1}{1 - \sum\limits_h \frac{\partial \mathbf{x}^h_d}{\partial E}}$ is the environmental feedback parameter.

The environmental feedback parameter σ and the sum of MWP^h_{EB} are positive as before, and the sign of the tax revenue term depends on how externality affects the demand. The two terms in the middle are again the interesting ones, because here the differences in the marginal willingnesses to pay (MWP^h_{EB}) refer to the redistribution question.

A type 2 household with higher productivity mimicking type 1 household gets more leisure, as they can do the work of type 1 faster and we have $\widehat{MWP}^2_{EB} > MWP^1_{EB}$. Thus the effect of the second term in the brackets in Eq. (13) is negative when environmental quality is a complement to leisure.

[11] For the derivation, see Appendix B.

This means that lowering the level of the externality increases type 2 household's incentive to mimic. To deter this, the government needs to redefine the income tax scheme by lowering taxes for type 2 and increasing taxes for type 1, i.e. let the income differences between type 1 and 2 households increase.

The third term refers to the difference $\widehat{MWP}_{EB}^{3} - MWP_{EB}^{2}$. Households 2 and 3 have the same productivities so the same conclusion as in the previous case does not hold. The difference can be solved by thinking about the deviation in preferences. When environmental quality, i.e. the negative of the environmental externality, is complement to leisure, type 3 households prefer leisure less and hence they do not "care" about the environmental quality as much as type 2 with their higher preference for leisure. This means that type 3 households are willing to pay less to avoid the externality than type 2 households, i.e. $\widehat{MWP}_{EB}^{3} < MWP_{EB}^{2}$ and the effect of the third term in Eq. (13) is positive. Thus, a more equal income distribution between type 2 and 3 households decreases the effect of this term by making mimicking less attractive and lowers the valuation of the harmfulness of the externality. As a result, the following proposition holds.

Proposition 1 *When environmental quality and leisure are complements, the valuation of the harmfulness of the externality is decreased when income differences between households 2 and 3 decrease and when income differences between type 1 and 2 households increase.*

The result implies that if the harmfulness of the externality increases with leisure, redistribution from high productivity households towards low productivity households impairs environmental quality whereas redistribution from more working high productivity household towards less working high productivity household improves environmental situation. Now the contradiction between environmental and redistributional aspects found in the earlier literature is no longer clear, because of the effect of term $\widehat{MWP}_{EB}^{3} - MWP_{EB}^{2}$. It is possible that the effect of the latter mimicking term is great enough to compensate the negative effect of the first mimicking term so that the overall relation between the two aspects is positive. In that case environmental and redistributional aspects would be in accordance from government's point of view.

One possible interpretation for individualss with a strong preference for leisure could be "laziness" and for those with a low preference for leisure as "hard working", as in Cuff (2000). The value of the harmfulness of the externality would diminish most when redistribution is directed from hard working low productivity households and from hard working high productivity households to lazy high productivity households. However, this direction of redistribution may not be the policy most supported by the majority.

It is also worth noting that, both in the separating and pooling equilibrium, if the preferences were separable between leisure and environmental quality and between leisure and consumption, the differences in MWP_{EB}^{h}, σ and the tax revenue term would all be zero. In the case of separable preferences the terms with the differences in MWP_{EB}^{h} will remain in the shadow price only when households have heterogeneous preferences with direct respect of environmental quality. However, here we have assumed that the preference for environmental quality is dependent on leisure and that the preferences are non-separable.

5 Commodity taxation

In the presence of an externality commodity taxes aim at internalising the harmful effect. Sandmo's additivity property (Sandmo, 1975) and Dixit's principle of targeting (1985), referred to here as the Sandmo-Dixit principle, state that the externality internalising part of the commodity tax should be separable from the other part of the tax rate and that it should affect only the tax rate of that good which creates the externality. There has been some discussion about the generality of the Sandmo-Dixit principle (Kopczuk, 2003). Here we study how an additional dimension in heterogeneity affects the principle.

In our framework the internalisation of the externality is not the only objective of taxation. Commodity taxation may get a new role as a screening tool in a framework with heterogeneous preferences. Blomquist and Christiansen (2004) report a result that in an economy with one consumption good and no externalities commodity taxation has an effect on redistribution and thus it can be used to mitigate the self-selection constraint.

The optimal commodity tax rates[12] are given by

$$\mathbf{t} = \frac{\Psi_{B^2}}{\gamma}\mathbf{S}^{+}\left(\mathbf{X}^2 - \mathbf{X}^1\right) - \frac{\lambda \widehat{V}_B^3}{\gamma}\mathbf{S}^{+}\left(\widehat{\mathbf{X}}^3 - \mathbf{X}^1\right) + \frac{\mu}{\gamma}\begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{14}$$

where, $\mathbf{S}^{+}$ is a unique pseudoinverse[13] of $\sum_h \mathbf{S}^{hT}$.

This form corresponds to that obtained in Blomquist and Christiansen (2004) with the exception of the last term here referring to externality. The second term implies that if a mimicker consumes more goods, i.e. if leisure and consumption are complements, $\widehat{\mathbf{X}}^3 > \mathbf{X}^1$ and commodity tax can be used to deter mimicking. Increasing commodity taxes and decreasing income taxes of type 1 households leaves a mimicker worse off and thereby making mimicking less attractive. If the difference in consumption is the other way around, i.e. $\widehat{\mathbf{X}}^3 < \mathbf{X}^1$, a negative commodity tax has the same effect.

The first term in Eq. (14) is interesting for two reasons. First it implies that if there is a difference in consumption between households of types 1 and 2, commodity taxation may also be used for redistribution. If type 2 households consume more, an increase in the commodity tax makes them worse off. In the opposite case a subsidy on commodities can be used. In the pooling equilibrium type 1 and 2 households were not observable, but now commodity taxes can be used as a screening device as it treats households differently.

The other important feature in the redistribution term is in the coefficient $\frac{\Psi_{B2}}{\gamma}$. It includes the shadow price of the externality $\frac{\mu}{\gamma}$. This means that when types 1 and 2 have nonidentical demand for goods, the externality effect appears in the tax rates of both commodities, also in the tax rate of the clean good. Before we can be sure of the failure of the Sandmo-Dixit principle, it is worth noting that when the shadow price $\frac{\mu}{\gamma}$ from (12) is substituted into (14), there is a possibility that the coefficient for $\frac{\Psi_{B2}}{\gamma}$ will cancel out to zero. The

[12] For the derivation, see Appendix C.

[13] To ensure the existence of the inverse, we use here pseudoinverse $\mathbf{S}^{+}$. Actually the Slutsky substitution matrix $\mathbf{S}$ probably also has the ordinary inverse $\mathbf{S}^{-1}$, as it is negative semidefinite. However, a pseudoinverse also exists for singular and non-square matrices and is equal to the ordinary inverse in the case of non-singular square matrix.

condition for this to happen is

$$\left(\mathbf{X}^{2T}-\mathbf{X}^{1T}\right)+\sigma\left(MWP_{EB}^{2}-MWP_{EB}^{1}\right)\sum_{h}\mathbf{s}_{d}^{h}=0\,. \tag{15}$$

Without more precise functional forms for the model we cannot rule out the possibility of this term being zero. In the special case where the condition holds, the first term on the right hand side of Eq. (14) cancels out and the shadow price reduces to

$$\frac{\mu}{\gamma}=\sigma\left\{\sum_{h}MWP_{EB}^{h}-\lambda*\left[\widehat{MWP}_{EB}^{3}-MWP_{EB}^{1}\right]-\sum_{h}\mathbf{t}^{T}\frac{\partial\mathbf{x}^{h}}{\partial E}\right\} \tag{16}$$

However, in a general case there is no need for the condition (15) to hold. Thus we can assume that generally term $\frac{\Psi_{B2}}{\gamma}$ remains in the optimal commodity tax rule and the Sandmo-Dixit principle fails in this two dimensional case in the pooling optimum. The following proposition summarises the result.

Proposition 2 *The effect of the externality on commodity taxes can no longer be separated to affect only the good that creates the externality. Thus the Sandmo-Dixit principle fails in the pooling optimum unless the consumption of the clean good is equal for pooling households.*

In the separating optimum the optimal commodity tax rule[14] corresponds those obtained in the earlier literature (Blomquist and Christiansen, 2004). The Sandmo-Dixit principle continues to hold and the ability to use commodity taxes to redistributional aims depends on the consumption behaviour of the mimickers. Under some assumptions (commodities are complements with leisure and substitute for each other and S is non-singular) increasing the tax on commodities mitigates one self-selection constraint and tightens the other. Thus using commodity taxes to mitigate the self-selection constraints includes a trade-off: one of the constraints can be relaxed at the expense of the other and the effect of commodity taxation on mimicking terms remains ambiguous.

Our analysis suggests that assuming a three-type economy is not sufficient to make the principle fail, but in the pooling equilibrium the externality based

[14] For details see Appendix C.

part can no longer be separated from the rest of the tax rate. In the separating case Ψ_{B^2} is actually a first order condition which requires that the marginal valuations for each group's hypothetical increase in income are zero, whereas in the pooling case this term is not (necessarily) zero. When the government wishes to redistribute away from type 2 i.e. $\Psi_{B^2} < 0$, in the pooling case commodity taxation is the only tax instrument that separates type 2 households, as long as the pooled households choose nonidentical consumption bundles. As the valuation of the hypothetical income depends on the externality created by an increased consumption of a dirty good, the optimal commodity tax of the clean good is also affected by the externality.

One explanation for the failure of the Sandmo-Dixit principle in the pooling case is the insufficient number of policy instruments. In the separating case income taxation takes care of the redistributional aims and commodity taxation internalises the externality. However, in the pooling case income taxes are not sufficient to handle redistribution, and the commodity tax has two policy objectives: redistribution and internalising the externality.

6 Conclusions

This study analyses the effect of a harmful environmental externality on the optimal tax policy in an economy with two-dimensional heterogeneity. We have assumed three types of households that differ both with respect to their productivities and their preferences for leisure. The valuation of the externality and the optimal commodity tax rates are discussed in two cases: in the pooling optimum, where the low productivity household and the high productivity household with a strong preference for leisure are assumed to choose the same income-consumption bundle, and in the separating optimum where each household chooses a different point.

The harmfulness of the externality measured by its shadow price is of the same form as in an economy with only one-dimensional heterogeneity. The term referring to mimicking suggests that there may be problems in combining environmental and redistributional preferences. The other term originating from the difference in preference for environmental quality may affect in either direction depending on the sign of its coefficient. However, with some assumptions of the desired direction of redistribution this term also indicates a contradiction between environmental and redistributional preferences. In the separating optimum one of the self-selection terms has a negative effect implying problems in combining environmental and redistributional aspects, as in the earlier literature. The other self-selection term has a positive effect. The valuation of the externality is decreased when income differences between high productivity households decrease and differences between high and low productivity households increase. Thus the contradiction observed in the earlier case is no longer clear.

The optimal commodity taxes in the pooling equilibrium offer two important results. The first is that commodity taxation can be used as a tool to differentiate and redistribute income between the households behaving identically as long as these households have unequal consumption of goods. If the household from which we want to distribute consumes more (less), a positive (negative) commodity tax makes them worse off and mitigates the self-selection constraint.

Another interesting question is the effect of the externality on commodity taxes. In the pooling optimum the Sandmo-Dixit principle fails to hold, i.e. the externality based part of the commodity tax can no longer be separated from the other part of the tax and it also affects the tax rate of the good not creating the harm. In the separating optimum we can generalise the Sandmo-Dixit principle. The explanation for this is that whereas in the separating case optimisation is done with respect to all types, in the pooling case the two household types are indistinguishable and the optimum is achieved with respect to two groups of households only. There are too few policy instruments in the pooling equilibrium: commodity taxes should take care of both redistribution and externality internalisation.

While taxes as a policy instrument to affect consumer behaviour are apparent, there are also other possibilities. As considered widely in environmental economics, instruments like regulations, quotas or trading permits are alternatives to environmental taxes. These instruments, however, are probably more suitable for affecting the production side of the economy. Thus, as a subject for future research it would be interesting to extend the analysis to a general equilibrium framework and compare the different policy instruments.

References

Beaudry, P. and C. Blackorby (1998), “Taxes and Employment Subsidies in Optimal Redistribution Programs”, NBER Working Paper No. 6355.

Aronsson, T. (2005), "Environmental Policy, Efficient Taxation and Unemployment", *International Tax and Public Finance* 12, 131-144.

Besley, T. and S. Coate (1995), “The Design of Income Maintenance Programmes”, *Review of Economic Studies* 62, 187-221.

Blomquist, S. and V. Christiansen (2004), “Taxation and Heterogeneous Preferences”, CESifo Working Paper No.1244.

Boadway, R., M. Marchand, P. Pestieau and M. del mar Racionero (2002), “Optimal Redistribution with Heterogeneous Preferences for Leisure”, *Journal of Public Economic Theory* 4, 475-498.

Bovenberg, A. L. and L. H. Goulder (2002), "Environmental Taxation and Regulation" in: A. J. Auerbach and M. Feldstein (eds.) "*Handbook of Public Economics, vol 3*", 1474-1540. North-Holland.

Cremer, H., F. Gahvari and N. Ladoux (1998), "Externalities and Optimal Taxation", *Journal of Public Economics* 70, 343-364.

Cremer, H., F. Gahvari and N. Ladoux (2001), "Second-Best Pollution Taxes and the Structure of Preferences", *Southern Economic Journal* 68, 258-280.

Cuff, K. (2000), “Optimality of workfare with Heterogeneous Preferences”, *Canadian Journal of Economics* 33, No. 1, 149-174.

Dixit, A. (1985), “A Tax Policy in Open Economies” in: A.J. Auerbach and M. Feldstein (eds.) “*Handbook of Public Economics, vol. 1*”, 313-374. North-Holland.

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

Fleurbaey, M. and F. Maniquet (1999), "Compensation and Responsibility", Mimeo, revised in 2005. Forthcoming in Arrow, K. J., A. K. Sen and K. Suzumura (eds.) "*Handbook of Social Choice and Welfare, Vol 2*". North-Holland.

Harrison, D. Jr. (1994), "*The Distributive Effects of Economic Instruments for Environmental Policy*", Paris: OECD.

Jacobsen, H. K., K. Birr-Pedersen and M. Wier (2003), "Distributional Implications of Environmental Taxation in Denmark". *Fiscal Studies* 24, No. 4, 477–499.

Jordahl, H. and L. Micheletto (2005), "Optimal Utilitarian Taxation and Horizontal Equity", *Journal of Public Economic Theory* 7, 681-708.

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

Mirrlees, J. A. (1976), "Optimal Tax Theory: A Synthesis", *Journal of Public Economics* 20, 77-87.

OECD (2001), "*Environmentally Related Taxes in OECD Countries, Issues and Strategies*", Paris.

Pirttilä, J. and M. Tuomala (1997), "Income Tax, Commodity Tax and Environmental Policy", *International Tax and Public Finance* 4, 379-393.

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. (1975), "Optimal Taxation in the Presence of Externalities", *Swedish Journal of Economics* 77, 86-98.

Sandmo, A. (1980), "Anomaly and Stability in the Theory of Externalities", *Quarterly Journal of Economics* 94, 799-807.

Sandmo, A. (1993), "Optimal Redistribution when Tastes Differ", *FinanzArchiv* 50, 149-163.

Smith, S. (1992), “Taxation and the Environment: A Survey”, *Fiscal Studies* 13, 21-57.

Stern, N. H. (1982), “Optimum 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.

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

Walls, M. and J. Hanson (1999), "Distributional aspects of an environmental tax shift: the case of motor vehicle emissions taxes", *National Tax Journal* vol. 52, 53–65.

Appendices

A Optimisation problem in the separating case

As a result of the similarity, the optimisation problem and the first order conditions of the separating optimum are presented only here. The Lagrange function of the optimisation problem is given by

$$
\begin{aligned}
L = {} & V^1\left(\mathbf{q}, B^1, L^1, E\right) \\
& + \delta_2\left[V^2\left(\mathbf{q}, B^2, L^2, E\right) - \overline{V}^2\right] + \delta_3\left[V^3\left(\mathbf{q}, B^3, L^3, E\right) - \overline{V}^3\right] \\
& + \lambda_2\left[V^2\left(\mathbf{q}, B^2, L^2, E\right) - V^2\left(\mathbf{q}, B^1, \Omega L^1, E\right)\right] \\
& + \lambda_3\left[V^3\left(\mathbf{q}, B^3, L^3, E\right) - V^3\left(\mathbf{q}, B^2, L^2, E\right)\right] \\
& + \gamma\left[\sum_h w^h L^h - \sum_h \mathbf{p}^T \mathbf{X}^h\left(\mathbf{q}, B^h, L^h, E\right) - \bar{G}\right] \\
& + \mu\left[E - \sum_h X_d^h\left(\mathbf{q}, B^h, L^h, E\right)\right]
\end{aligned} \tag{17}
$$

The corresponding first order conditions with respect to L^h, B^h for $h = 1, 3$, E and $\mathbf{q}$ are[15]

$$V_L^1 - \lambda_2 \widehat{V}_L^2 \Omega + \gamma\left(w^1 - \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial L^1}\right) - \mu \frac{\partial X_d^1}{\partial L^1} = 0 \tag{18}$$

$$V_B^1 - \lambda_2 \widehat{V}_B^2 - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial B^1} - \mu \frac{\partial X_d^1}{\partial B^1} = 0 \tag{19}$$

$$\left(\delta_2 + \lambda_2\right) V_L^2 - \lambda_3 \widehat{V}_L^3 + \gamma\left(w^2 - \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial L^2}\right) - \mu \frac{\partial X_d^2}{\partial L^2} = 0 \tag{20}$$

[15] Note that $\mathbf{X}^h$ is a function of L^h and B^h only, thus sums can be dropped out in the first six derivatives.

$$(\delta_2+\lambda_2)\,V_B^2-\lambda_3\widehat{V}_B^3-\gamma\mathbf{p}^T\frac{\partial\mathbf{X}^2}{\partial B^2}-\mu\frac{\partial X_d^2}{\partial B^2}=0 \tag{21}$$

$$(\delta_3+\lambda_3)\,V_L^3+\gamma\left(w^3-\mathbf{p}^T\frac{\partial\mathbf{X}^3}{\partial L^3}\right)-\mu\frac{\partial X_d^3}{\partial L^3}=0 \tag{22}$$

$$(\delta_3+\lambda_3)\,V_B^3-\gamma\mathbf{p}^T\frac{\partial\mathbf{X}^3}{\partial B^3}-\mu\frac{\partial X_d^3}{\partial B^3}=0 \tag{23}$$

$$\begin{aligned}V_E^1+(\delta_2+\lambda_2)\,V_E^2+(\delta_3+\lambda_3)\,V_E^3-\lambda_2\widehat{V}_E^2-\lambda_3\widehat{V}_E^3&\\-\gamma\sum_h\mathbf{p}^T\frac{\partial\mathbf{X}^h}{\partial E}+\mu\left(1-\sum_h\frac{\partial X_d^h}{\partial E}\right)&=0\end{aligned} \tag{24}$$

$$\begin{aligned}V_{\mathbf{q}}^1+(\delta_2+\lambda_2)\,V_{\mathbf{q}}^2+(\delta_3+\lambda_3)\,V_{\mathbf{q}}^3-\lambda_2\widehat{V}_{\mathbf{q}}^2-\lambda_3\widehat{V}_{\mathbf{q}}^3&\\-\gamma\sum_h\mathbf{p}^T\frac{\partial\mathbf{X}^h}{\partial\mathbf{q}}-\mu\sum_h\frac{\partial X_d^h}{\partial\mathbf{q}}&=0\,,\end{aligned} \tag{25}$$

where again the terms with a 'hat' refer to mimickers.

B The valuation of externality

In the derivation we use the compensated conditional demand function, given by minimising households' expenditure conditional on keeping the direct utility U on a given level u, i.e.

$$\begin{aligned}\mathbf{x}^h\left(\mathbf{q},Y^h,u^h,E,w^h\right)&=\\&\arg\min_{\mathbf{X}^h}\left\{\mathbf{q}^T\mathbf{X}^h\left(\mathbf{q},B^h,L^h,E\right)\left|U\left(X^h,Y^h/w^h,E\right)\geqslant u^h\right.\right\}\end{aligned} \tag{26}$$

As is shown in Edwards, Keen and Tuomala (1994), the following Slutsky-type properties hold[16]:

[16] Although in Edwards et al. (1994) these are shown for a public good, the case of public

$$\sum_h \mathbf{p}^T \frac{\partial \mathbf{x}^h}{\partial E} = \sum_h MWP_{EB}^h - \sum_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E} \tag{27}$$

$$\frac{\partial \mathbf{X}^h}{\partial E} = \frac{\partial \mathbf{x}^h}{\partial E} - MWP_{EB}^h \frac{\partial \mathbf{X}^h}{\partial B^h} \tag{28}$$

from which is follows that

$$\sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E} = \sum_h \mathbf{p}^T \frac{\partial \mathbf{x}^h}{\partial E} - \sum_h MWP_{EB}^h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^h} \tag{29}$$

The pooling case

To derive the shadow price add and subtract terms $\lambda \hat{V}_B^3 \frac{V_E^1}{V_B^1}$ and $\delta_2 V_B^2 \frac{V_E^1}{V_B^1}$ from Equation (10) and denote $-\frac{V_E^h}{V_B^h} = MWP_{EB}^h$ to get

$$\begin{aligned}\left[-V_B^1 - \delta_2 V_B^2 + \lambda \widehat{V}_B^3\right] MWP_{EB}^1 - \delta_2 V_B^2 \left[MWP_{EB}^2 - MWP_{EB}^1\right] \\ - (\delta_3 + \lambda) V_B^3 MWP_{EB}^3 + \lambda \widehat{V}_B^3 \left[\widehat{MWP}_{EB}^3 - MWP_{EB}^1\right] \\ - \gamma \sum_{h=1,2,3} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E} + \mu \left(1 - \sum_{h=1,2,3} \frac{\partial X_d^h}{\partial E}\right) = 0\end{aligned} \tag{30}$$

Next use the first order conditions (7): $-V_B^1 - \delta_2 V_B^2 + \lambda \widehat{V}_B^3 = -\gamma \sum_{h=1,2} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^1} - \mu \sum_{h=1,2} \frac{\partial X_d^h}{\partial B^1}$ and (9) : $-(\delta_3 + \lambda) V_B^3 = -\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^3}{\partial B^3} - \mu \frac{\partial X_d^3}{\partial B^3}$ to rewrite this as

$$\begin{aligned}- \left[\gamma \sum_{h=1,2} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^1} + \mu \sum_{h=1,2} \frac{\partial X_d^h}{\partial B^1}\right] MWP_{EB}^1 - \delta_2 V_B^2 \left[MWP_{EB}^2 - MWP_{EB}^1\right] \\ - \left[\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^3}{\partial B^3} + \mu \frac{\partial X_d^3}{\partial B^3}\right] MWP_{EB}^3 + \lambda_3 \widehat{V}_B^3 \left[\widehat{MWP}_{EB}^3 - MWP_{EB}^1\right] \\ - \gamma \sum_{h=1,2,3} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E} + \mu \left(1 - \sum_{h=1,2,3} \frac{\partial X_d^h}{\partial E}\right) = 0\end{aligned} \tag{31}$$

---

bad considered here goes analogously.

Using properties (27) and (29) subsequently and dividing by γ we obtain

$$
\begin{aligned}
& - MWP^1_{EB}\left[\mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^1} + \frac{\mu}{\gamma}\frac{\partial X^2_d}{\partial B^1}\right] + MWP^2_{EB}\left[\mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2} + \frac{\mu}{\gamma}\frac{\partial X^2_d}{\partial B^2}\right] \\
& \quad + \frac{\lambda \hat{V}^3_B}{\gamma}\left[M\hat{W}P^3_{EB} - MWP^1_{EB}\right] + \frac{\delta_2 V^2_B}{\gamma}\left[MWP^1_{EB} - MWP^2_{EB}\right] \\
& \quad - \left[\sum_{h=1,2,3} MWP^h_{EB} - \sum_{h=1,2,3} \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E}\right] + \frac{\mu}{\gamma}\left(1 - \sum_{h=1,2,3}\frac{\partial x^h_d}{\partial E}\right) = 0 \quad (32)
\end{aligned}
$$

Because $B^1 = B^2$, it must be that $\frac{\partial X^2}{\partial B^1} = \frac{\partial X^2}{\partial B^2}$. Thus we obtain

$$
\begin{aligned}
& \left[\frac{\delta_2 V^2_B}{\gamma} - \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^1} - \frac{\mu}{\gamma}\frac{\partial X^2_d}{\partial B^1}\right]\left[MWP^1_{EB} - MWP^2_{EB}\right] \\
& \quad + \frac{\lambda \hat{V}^3_B}{\gamma}\left[M\hat{W}P^3_{EB} - MWP^1_{EB}\right] - \left[\sum_{h=1,2,3} MWP^h_{EB} - \sum_{h=1,2,3} \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E}\right] \\
& \qquad + \frac{\mu}{\gamma}\left(1 - \sum_{h=1,2,3}\frac{\partial x^h_d}{\partial E}\right) = 0 \quad (33)
\end{aligned}
$$

Denoting $-\frac{\delta_2 V^2_B}{\gamma} + \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2} + \frac{\mu}{\gamma}\frac{\partial X^2_d}{\partial B^2} = \frac{\Psi_{B^2}}{\gamma}$ and $\frac{\lambda_3 \widehat{V}^3_B}{\gamma} = \lambda^*$, we can solve $\frac{\mu}{\gamma}$ to get the implicit solution for the shadow price in (12). The explicit solution in footnote 7 can be found with the help of Eq (33) by leaving all terms with the shadow price $\frac{\mu}{\gamma}$ on the left hand side and moving everything else to the right hand side.

The separating case

Letting $-\frac{V_E^h}{V_B^h} = MWP_{EB}^h$ denote the marginal willingness to pay to avoid the externality, we can write Equation (24) as

$$
\begin{aligned}
&-\left[V_B^1 - \lambda_2 \widehat{V}_B^2\right] MWP_{EB}^1 - \left[(\delta_2 + \lambda_2) V_B^2 - \lambda_3 \hat{V}_B^3\right] MWP_{EB}^2 \\
&\qquad - (\delta_3 + \lambda_3) V_B^3 MWP_{EB}^3 + \lambda_2 \widehat{V}_B^2 \left[\widehat{MWP}_{EB}^2 - MWP_{EB}^1\right] \\
&+ \lambda_3 \widehat{V}_B^3 \left[\widehat{MWP}_{EB}^3 - MWP_{EB}^2\right] - \gamma \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E} + \mu \left(1 - \sum_h \frac{\partial X_d^h}{\partial E}\right) = 0
\end{aligned} \tag{34}
$$

Using the first order conditions in Equations (19), (21) and (23) this can be written as

$$
\begin{aligned}
&- MWP_{EB}^1 \left[\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial B^1} + \mu \frac{\partial X_d^1}{\partial B^1}\right] - MWP_{EB}^2 \left[\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2} + \mu \frac{\partial X_d^2}{\partial B^2}\right] \\
&\qquad - MWP_{EB}^3 \left[\gamma \mathbf{p}^T \frac{\partial \mathbf{X}^3}{\partial B^3} + \mu \frac{\partial X_d^3}{\partial B^3}\right] + \lambda_2 \widehat{V}_B^2 \left[\widehat{MWP}_{EB}^2 - MWP_{EB}^2\right] \\
&+ \lambda_3 \widehat{V}_B^3 \left[\widehat{MWP}_{EB}^3 - MWP_{EB}^2\right] - \gamma \sum_h \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial E} + \mu \left(1 - \sum_h \frac{\partial X_d^h}{\partial E}\right) = 0
\end{aligned} \tag{35}
$$

Using properties (27) and (29), and dividing both sides by γ and rearranging the terms we get

$$
\begin{aligned}
&\frac{\lambda_2 \widehat{V}_B^2}{\gamma} \left[\widehat{MWP}_{EB}^2 - MWP_{EB}^1\right] + \frac{\lambda_3 \widehat{V}_B^3}{\gamma} \left[\widehat{MWP}_{EB}^3 - MWP_{EB}^2\right] \\
&\qquad - \sum_h MWP_{EB}^h + \sum_h \mathbf{t}^T \frac{\partial \mathbf{x}^h}{\partial E} + \frac{\mu}{\gamma} \left(1 - \sum_h \frac{\partial x_d^h}{\partial E}\right) = 0
\end{aligned} \tag{36}
$$

Simplifying the equation by cancelling out terms, denoting $\frac{\lambda_2 \widehat{V}_B^2}{\gamma} = \lambda_2^*$ and $\frac{\lambda_3 \widehat{V}_B^3}{\gamma} = \lambda_3^*$ and multiplying by $\left(1 - \sum_h \frac{\partial x_d^h}{\partial E}\right)^{-1} =: \sigma$ gives (13).

C The commodity tax rules

Derivation in the pooling case

Using Slutsky decomposition $\frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} = \mathbf{S}^h - \frac{\partial \mathbf{X}^h}{\partial B^h}\mathbf{X}^{hT}$, where $\mathbf{S}^h = \begin{bmatrix} s^h_{cc} & s^h_{cd} \\ s^h_{dc} & s^h_{dd} \end{bmatrix} = \begin{bmatrix} \mathbf{s}^h_c \\ \mathbf{s}^h_d \end{bmatrix}$, and Roy's identity $V^h_{\mathbf{q}} = -V^h_B \mathbf{X}^{hT}$, the first order condition (11) can be written as

$$\begin{aligned} &- V^1_B \mathbf{X}^{1T} - \delta_2 V^2_B \mathbf{X}^{2T} - (\delta_3 + \lambda) V^3_B \mathbf{X}^{3T} + \lambda \widehat{V}^3_B \widehat{\mathbf{X}}^{3T} \\ &\quad - \gamma \sum_{h=1,2,3} \mathbf{p}^T \left(\mathbf{S}^h - \frac{\partial \mathbf{X}^h}{\partial B^h} \mathbf{X}^{hT} \right) - \mu \sum_{h=1,2,3} \left(\mathbf{s}^h_d - \frac{\partial X^h_d}{\partial B^h} \mathbf{X}^{hT} \right) = 0 \end{aligned} \tag{37}$$

Substitute the first order condition (7) multiplied (from the right) by $\mathbf{X}^{1T}$ and the first order condition (9) multiplied by $\mathbf{X}^{3T}$ to get

$$\begin{aligned} &- \delta_2 V^2_B \left(\mathbf{X}^{2T} - \mathbf{X}^{1T} \right) + \lambda \widehat{V}^3_B \left(\widehat{\mathbf{X}}^{3T} - \mathbf{X}^{1T} \right) - \gamma \sum_{h=1,2} \mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^1} \mathbf{X}^{1T} \\ &\quad - \mu \sum_{h=1,2} \frac{\partial X^h_d}{\partial B^1} \mathbf{X}^{1T} - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^3}{\partial B^3} \mathbf{X}^{3T} - \mu \frac{\partial X^3_d}{\partial B^3} \mathbf{X}^{3T} \\ &\quad - \gamma \sum_{h=1,2,3} \mathbf{p}^T \left(\mathbf{S}^h - \frac{\partial \mathbf{X}^h}{\partial B^h} \mathbf{X}^{hT} \right) - \mu \sum_{h=1,2,3} \left(\mathbf{s}^h_d - \frac{\partial X^h_d}{\partial B^h} \mathbf{X}^{hT} \right) = 0 \end{aligned} \tag{38}$$

Remembering that in the pooling equilibrium $B^1 = B^2$ and thus also $\frac{\partial X^h}{\partial B^1} = \frac{\partial X^h}{\partial B^2}$, and by reorganising the terms we can simplify the previous equation to the form

$$\left[-\delta_2 V_B^2 + \mu\frac{\partial X_d^2}{\partial B^2} + \gamma\mathbf{p}^T\frac{\partial \mathbf{X}^2}{\partial B^2}\right]\left(\mathbf{X}^{2T} - \mathbf{X}^{1T}\right)$$
$$+ \lambda\widehat{V}_B^3\left(\widehat{\mathbf{X}}^{3T} - \mathbf{X}^{1T}\right) - \gamma\sum_{h=1,2,3}\mathbf{p}^T\mathbf{S}^h - \mu\sum_{h=1,2,3}\mathbf{s}_d^h = 0 \quad (39)$$

With the help of a differential from the consumer's budget constraint it can be noted that $\sum_h \mathbf{q}^T\mathbf{S}^h = 0$.[17] Note also that $\sum_h s_d^h = \begin{bmatrix} 0 & 1 \end{bmatrix}\sum_h \mathbf{S}^h$. Making use of the definition of $\Psi_{B^2} = \delta_2 V_B^2 - \gamma\mathbf{p}^T\frac{\partial \mathbf{X}^2}{\partial B^2} - \mu\frac{\partial X_d^2}{\partial B^2}$, reorganising terms and dividing them by γ yields

$$\sum_{h=1,2,3}\mathbf{t}^T\mathbf{S}^h = \frac{\Psi_{B^2}}{\gamma}\left(\mathbf{X}^{2T} - \mathbf{X}^{1T}\right) - \frac{\lambda\widehat{V}_B^3}{\gamma}\left(\widehat{\mathbf{X}}^{3T} - \mathbf{X}^{1T}\right) + \frac{\mu}{\gamma}\sum_{h=1,2,3}\mathbf{s}_d^h \quad (40)$$

Finally, taking a transpose, multiplying this from the left by $\mathbf{S}^+$, a unique pseudoinverse (the Moore-Penrose inverse) of $\sum_h \mathbf{S}^{hT}$, gives us optimal commodity tax rates in (14).

Commodity taxes in the separating optimum

Using Roy's identity $V_\mathbf{q}^h = -V_B^h\mathbf{X}^{hT}$ and Slutsky decomposition $\frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} = \mathbf{S}^h - \frac{\partial \mathbf{X}^h}{\partial B^h}\mathbf{X}^{hT}$ the first order condition (25) can be written as

$$- V_B^1\mathbf{X}^{1T} - (\delta_2 + \lambda_2)\, V_B^2\mathbf{X}^{2T} - (\delta_3 + \lambda_3)\, V_B^3\mathbf{X}^{3T} + \lambda_2\widehat{V}_B^2\widehat{\mathbf{X}}^{2T} + \lambda_3\widehat{V}_B^3\widehat{\mathbf{X}}^{3T}$$
$$- \gamma\sum_{h=1,2,3}\mathbf{p}^T\left(\mathbf{S}^h - \frac{\partial \mathbf{X}^h}{\partial B^h}\mathbf{X}^{hT}\right) - \mu\sum_{h=1,2,3}\left(\mathbf{s}_d^h - \frac{\partial X_d^h}{\partial B^h}\mathbf{X}^{hT}\right) = 0\,, \quad (41)$$

[17] Differentiating consumer budget constraint $\mathbf{q}^T\mathbf{X} = B$ gives $\mathbf{q}^T\frac{\partial \mathbf{X}}{\partial B} = 1$ and $\left[\frac{\partial \mathbf{X}}{\partial \mathbf{q}}\right]^T\mathbf{q} = -\mathbf{X}$. Substituting these into $\sum_h \mathbf{q}^T\mathbf{S}^h = \sum_h \mathbf{q}^T\left(\frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} + \frac{\partial \mathbf{X}^h}{\partial B^h}\mathbf{X}^{hT}\right)$ yields $\sum_h \mathbf{q}^T\mathbf{S}^h = \sum_h \mathbf{q}^T\frac{\partial \mathbf{X}^h}{\partial \mathbf{q}} + \sum_h \mathbf{q}^T\frac{\partial \mathbf{X}^h}{\partial B^h}\mathbf{X}^{hT} = -\sum_h \mathbf{X}^{hT} + \sum_h \mathbf{X}^{hT} = 0$

Substituting in the first order conditions (19) multiplied (from the right) by $\mathbf{X}^{1T}$, (21) multiplied by $\mathbf{X}^{2T}$ and (23) multiplied by $\mathbf{X}^{3T}$ gives

$$\begin{aligned}
&-\lambda_2 \widehat{V}_B^2 \mathbf{X}^{1T} - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^1}{\partial B^1}\mathbf{X}^{1T} - \mu \frac{\partial X_d^1}{\partial B^1}\mathbf{X}^{1T} \\
&\qquad - \lambda_3 \widehat{V}_B^3 \mathbf{X}^{2T} - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^2}{\partial B^2}\mathbf{X}^{2T} - \mu \frac{\partial X_d^2}{\partial B^2}\mathbf{X}^{2T} \\
&\qquad - \gamma \mathbf{p}^T \frac{\partial \mathbf{X}^3}{\partial B^3}\mathbf{X}^{3T} - \mu \frac{\partial X_d^3}{\partial B^3}\mathbf{X}^{3T} + \lambda_2 \widehat{V}_B^2 \widehat{\mathbf{X}}^{2T} + \lambda_3 \widehat{V}_B^3 \widehat{\mathbf{X}}^{3T} \\
&\qquad - \gamma \sum_{h=1,2,3} \mathbf{p}^T \left(\mathbf{S}^h - \frac{\partial \mathbf{X}^h}{\partial B^h}\mathbf{X}^{hT} \right) - \mu \sum_{h=1,2,3} \left(\mathbf{s}_d^h - \frac{\partial X_d^h}{\partial B^h}\mathbf{X}^{hT} \right) = 0\,, \qquad (42)
\end{aligned}$$

reorganising and cancelling out terms $\mathbf{p}^T \frac{\partial \mathbf{X}^h}{\partial B^h}\mathbf{X}^{hT}$ and $\frac{\partial X_d^h}{\partial B}\mathbf{X}^{hT}$ gives

$$\begin{aligned}
&\lambda_2 \widehat{V}_B^2 \left(\widehat{\mathbf{X}}^{2T} - \mathbf{X}^{1T} \right) \\
&\qquad + \lambda_3 \widehat{V}_B^3 \left(\widehat{\mathbf{X}}^{3T} - \mathbf{X}^{2T} \right) \\
&\qquad\qquad - \gamma \sum_{h=1,2,3} \mathbf{p}^T \mathbf{S}^h - \mu \sum_{h=1,2,3} \mathbf{s}_d^h = 0\,, \qquad (43)
\end{aligned}$$

using $\mathbf{p}^T = \mathbf{q}^T - \mathbf{t}^T$ and the result $\sum_h \mathbf{q}^T \mathbf{S}^h = 0$, dividing by γ and reorganising yields

$$\sum_{h=1,2,3} \mathbf{t}^T \mathbf{S}^h = -\frac{\lambda_2 \widehat{V}_B^2}{\gamma} \left(\widehat{\mathbf{X}}^{2T} - \mathbf{X}^{1T} \right) - \frac{\lambda_3 \widehat{V}_B^3}{\gamma} \left(\widehat{\mathbf{X}}^{3T} - \mathbf{X}^{2T} \right) + \frac{\mu}{\gamma} \sum_{h=1,2,3} \mathbf{s}_d^h \quad (44)$$

And finally, taking a transpose, multiplying (from the left) by $\mathbf{S}^+$, a unique pseudoinverse (the Moore-Penrose inverse) of $\sum_h \mathbf{S}^{hT}$, yields

$$\mathbf{t} = -\lambda_2^* \mathbf{S}^+ \left(\widehat{\mathbf{X}}^2 - \mathbf{X}^1 \right) - \lambda_3^* \mathbf{S}^+ \left(\widehat{\mathbf{X}}^3 - \mathbf{X}^2 \right) + \frac{\mu}{\gamma} \begin{bmatrix} 0 \\ 1 \end{bmatrix} . \qquad (45)$$

where $\lambda_2^* = \frac{\lambda_2 \widehat{V}_B^2}{\gamma}$ and $\lambda_3^* = \frac{\lambda_3 \widehat{V}_B^3}{\gamma}$.

The first thing to notice is the last term including the effect of externality. As the valuation of the externality, $\frac{\mu}{\gamma}$, does not appear in $\mathbf{S}^+$, it only affects the tax rate of the dirty good $\mathbf{X}_d$. Thus the externality based part of the commodity tax can be separated and only affects the tax rate of the good creating the externality, i.e. the Sandmo-Dixit principle continues to hold in this framework.

If we assume that commodities $\mathbf{X}_c$ and $\mathbf{X}_d$ are substitutes and matrix $\mathbf{S}$ is nonsingular, we find that all elements in $\mathbf{S}^+$ are negative[18]. In that case the two first terms are positive if mimickers consume more than households revealing their true types. This implies that commodity taxation can be used to deter mimicking. In this model types 2 and 3 have equal productivities and furthermore type 3 households prefer leisure less than type 2 households. If the demand for goods increases with leisure and they are taxed, type 2 households find mimicking low productivity households less attractive and the self-selection constraint between these households is mitigated. Type 3 mimickers have exactly the same amount of leisure as true type 2 households and thus commodity taxation has no effect on mimicking via leisure. However, when consumption is complement with leisure, type 3 mimickers with less preference for leisure also prefer less consumption. Now type 2 households would suffer more from increased commodity tax than type 3 mimickers and the welfare decreasing effect of the self-selection constraint become worse. Thus using commodity taxes to deter mimicking includes a trade-off: increasing the tax mitigates one self-selection constraint and tightens the other.

[18] Without more precise functional forms we cannot determine the signs of the elements in $\mathbf{S}^+$ unambiguously.

Chapter 4

Pawns and Queens Revisited: Public Provision of Private Goods When Individuals Make Mistakes*

Jukka Pirttilä
Labour Institute for Economic Research
and
Sanna Tenhunen
University of Tampere and FDPE

Abstract: This paper analyses the optimal tax policy and public provision of private goods when the government is paternalistic and has a redistributive objective. When individuals only differ with respect to their income-earning abilities, the publicly-provided goods should be overprovided, relative to the decentralised optimum, if society's marginal valuation of them exceeds the individual valuation and if these goods are complements to labour supply. However, when the individuals also differ in terms of their valuation of the publicly-provided good, this simple conclusion does not hold. Optimal marginal income tax rates are shown to differ from the standard rules if publicly-provided goods and labour supply are related.

Keywords: Behavioural economics, optimal taxation, public provision.
JEL classification: H21, H42

*Acknowledgements: We are grateful to three anonymous referees and the editor, John D. Wilson, as well as Timo Seppälä, Matti Tuomala, and seminar participants at Uppsala University, FIEF, HECER, and CESifo Area Conference for Public Economics in 2005 for useful comments.

1 Introduction

Recent advances in behavioural economics have demonstrated that individual decision-making suffers from bounded rationality and various biases.[1] These biases can be especially problematic in complex decisions that involve uncertainty and dynamism. In these situations, the government might want to intervene, indeed individuals might want the government to intervene, to induce behaviour that is closer to what individual wish they were doing. The analysis of such corrective interventions, e.g. through taxes and subsidies, can be called 'behavioural public economics'. In these cases, where the government has an objective function that is different from that of individuals, the government is said to be 'non-welfarist' or paternalistic in its objectives.[2]

One important area where the government could improve upon individual choice is related to goods such as education, health and insurance, which in many countries are indeed often publicly regulated, provided or subsidized. In a recent book, Le Grand (2003) discusses whether the clients of government services should be treated as pawns (that is, recipients whose decisions are mainly delegated to the provider) or queens (sovereign consumers). There are a number of reasons why individuals are particularly prone to make mistakes, for instance, decisions related to health (which will be our emphasis below). First, the quantity of information may simply be too great or the causal connections too difficult to understand, relative to the mental capacity of a majority of individuals. Second, as in the case of education, society might want to make some of the decisions on behalf of the parents to protect the children's rights. Finally, returns to investments in health often accrue only in the distant future. If individuals have a tendency to undervalue future benefits (e.g. because of hyperbolic discounting, Laibson (1997)), they might be better off if they delegated some of the decision-making to an outsider, e.g. the government, to protect themselves against their own weakness of will. Treating the customers of public services as pawns instead of queens in certain situations can, therefore,

[1] For a survey, see for example Camerer and Loewenstein (2004).

[2] Recent examples of this rapidly expanding field of research include O'Donoghue and Rabin (2003), Sheshinski (2003), Gruber and Köszegi (2004) and Aronsson and Thurström (2005). See Seade (1980) for a seminal analysis and Glaeser (2005) for a critical view.

be desirable.

Our aim is to take these points seriously and consider the optimal public provision of private goods, when individuals' demand for these goods suffers from the sort of mistakes behavioural economics has highlighted. We examine what will happen to optimal policy if the government tries to correct these mistakes by basing its own decision on what it thinks is truly best for the individuals. In other words, the government's objective function is paternalistic or non-welfarist. Problems of self control could perhaps be corrected by market mechanisms, e.g. by supply of commitment devices. Köszegi (2005) offers a discussion of these options. He concludes that competition between providers of commitment devices limit the market's ability to offer the individuals ways to overcome weakness of will. Therefore, some form of public intervention may be necessary to induce individual choices closer to the "true" optimum.

The paper builds on what has become a now standard framework in the literature, that is, the analysis of public provision as a part of the government's redistributive system. In the modern information-based approach to tax analysis, initiated by Mirrlees (1971), there is, by now, a large literature examining the role of publicly-provided goods (e.g. Boadway and Keen 1993, Edwards, Keen and Tuomala 1994, Boadway and Marchand 1995, Cremer and Gahvari 1997, and Pirttilä and Tuomala 2002). Because of asymmetric information between the taxpayers and the government, the government must take individuals' incentives into account in its optimal tax policy (technically speaking, through incentive-compatibility constraints). Public provision can then be a useful tool for redistribution if it helps to relax the harmful incentive effects of income taxation.

The present paper differs from the existing literature by assuming that the government's objective function is non-welfaristic, but we still use the same workhorse model as our benchmark. We, therefore, account for two potential constraints to optimality: asymmetric information (as in earlier models) and mistakes in individual decision-making. It is interesting to examine precisely this framework to see how the two departures from the first best and, consequently, two motivations for public intervention interact. Do they lead to similar results or are there situations where the simultaneity of the two aims lead to ambiguous policy conclusions?

Instead of assuming that all individuals make similar mistakes, it is much more realistic to allow, following O'Donoghue and Rabin (2003), for differences in rationality.[3] We, therefore, examine the case where individuals differ in two respects, rationality and income-earning abilities (which is clearly needed for the redistribution motive to make sense). Deriving clear-cut results in optimal tax analysis when individuals differ in more than one respect is complicated (see, for example, the discussion in Boadway, Marchand, Pestieau and Racionero 2002). We, therefore, follow Blomquist and Christiansen (2003, 2004) and concentrate on a three-type interpretation of the Mirrlees (1971) model. Building on Stern (1982) and Stiglitz (1982), households can be divided into skilled and less-skilled groups. In addition, one of the groups can be either fully rational or partly irrational. Proceeding with the three-type case allows for a much easier intuitive discussion, and yet it fully captures the key mechanisms at work.

Our paper is most closely related to Kanbur, Pirttilä and Tuomala (2006) and Blomquist and Micheletto (2006),[4] who both examine non-welfarist income and commodity taxation. The former considers commodity taxation of merit goods as a part of optimal mixed taxation in non-welfarist framework: subsidization of such goods is found unambiguously meritorious. The latter focuses on a view that when there are several goods in the economy, it becomes unclear whether merit goods should be subsidized.[5] However, neither the public provision of private goods nor two-dimensional heterogeneity is considered in these papers.

While price subsidies and public provision can both be used as corrective

[3]In the paper, we refer to someone as 'irrational' when his or her preferences differ from what is traditionally seen as rational. Since the government is assumed to be purely rational, an irrational person can, thus, also be described as someone whose preferences differ from the government's preferences.

[4]We learnt about the work by Blomquist and Micheletto (2006) after having finished a first draft of our paper.

[5]In Blomquist and Micheletto (2006) the finding that subsidizing merit good is not necessarily desirable results from the generality of model: the optimality of subsidizing merit good depends e.g. on whether other goods are substitutes or complements to the merit good, whether mimicker's marginal rate of substitution is greater or smaller than true type's MRS and how strong policy maker's distributional concerns are.

devices, public provision can be especially important in situations with individual mistakes. With public provision, the government can make sure that the individuals consume a certain minimum amount of goods deemed meritorious. Examples of such goods include compulsory health insurance or education. One rationale for using public provision instead of subsidization is 'specific egalitarianism' suggesting that some goods and services should be distributed more equally than what market outcome would be.[6]

At a conceptual level, our study is also related to earlier work, beginning from Musgrave (1959), on merit goods. The optimal tax treatment of merit goods, but not their optimal public provision, is analysed by Sandmo (1983), Besley (1988), Racionero (2001), and Schroyen (2005).

The paper proceeds as follows. Section 2 presents a benchmark model where paternalistic concerns are introduced into the framework often used in the analysis of publicly-provided goods. Section 3 introduces our main idea, i.e. how differences in rationality complicate the design of optimal public provision. Section 4 discusses an application of the results to a concrete policy problem, decisions regarding health behaviour. Section 5 concludes.

2 The basic model: All irrational

Consider a Stiglitz (1982) -type model with public provision of private goods along the lines of Boadway and Marchand (1995).[7] There are two types of households, 1 and 2. The number of types in each group is $N^i (> 0)$, such that $\sum N^i = 1$. The wage rates of the households are $w_1 < w_2$. The households supply labour l, and their gross income is $y = wl$. The households' skill levels are private information, and the government must design a tax schedule based on observable income instead. The after-tax income of a household is given by $x = y - T(y)$, where $T(y)$ is a non-linear tax schedule set by the government. The household can spend its after-tax income on two goods, a normal consumption good, c, and on another normal good, e, which is also provided by the government. The extent of government provision is denoted

[6] The concept of specific egalitarianism was introduced by Tobin (1970)

[7] A similar framework is also used and thoroughly explained in Blomquist and Christiansen (1998a, 1998b).

by g; the overall amount available to the household is $e + g \equiv z$. In other words, the households can top up the publicly-provided good through their own purchases, e.[8] Households' utility is given by a strictly quasi-concave direct utility function $U\left(c, \frac{y}{w}, z\right)$, where U is increasing in both goods c and z and decreasing in hours of labour l $(= \frac{y}{w})$. Households' budget constraint is $c + e = x$, with prices of the consumption good c and the privately purchased part of good z normalised to one. Thus, we define households' utility function as

$$V^i\left(x^i - e^i, y^i, g + e^i\right) \equiv U\left(x^i - e^i, \frac{y^i}{w^i}, g + e^i\right) \quad (1)$$

The households maximise their utility given by $V^i\left(x^i - e^i, y^i, g + e^i\right)$ with respect to privately purchased part e. Since the privately purchased part cannot be negative, we have Kuhn-Tucker conditions

$$-V_c^i + V_e^i \leq 0 \text{ and } e^i\left(-V_c^i + V_e^i\right) = 0 \quad (2)$$

This first-stage optimization provides us the demand for e, and substituting it back to utility function $V^i\left(x^i - e^i, y^i, g + e^i\right)$ gives us the partially indirect utility function of the households, $v(x^i, y^i, g)$, consisting of arguments that are observable for the government: post tax income x, pre-tax income y to capture the effect of labour supply and publicly-provided part g.

Examples of the publicly-provided good can include investment in education or health services, social insurance, child care and care of the elderly. Stylized facts suggest that many of these goods are complements to labour supply: daycare and care of the elderly are clearly more important for workers than non-workers, and public education becomes useful when employed. Following much of the earlier work in this area, we therefore also assume throughout the paper that the publicly-provided good and labour supply are complements.[9]

The key point is that households can rationally and without biases select

[8] Note that, for simplicity, the household must indeed consume at least the amount equivalent to g; in other words, it is assumed that the government has means to control consumption. The case where control is costly is an important avenue of future research.

[9] The results in a case where publicly-provided good and labour supply are substitutes are presented in the working paper version (Pirttilä and Tenhunen, 2005).

the consumption of purely private goods, x.[10] However, they have difficulties in determining the correct level of investment on the publicly-provided goods. This may be due to the complexity of the decision. A particularly interesting case is one where the benefits of publicly-provided goods are delayed. If households use hyperbolic discounting, they tend to purchase too little of the publicly-provided good from the point of view of their period 0 selves and social welfare. Consider, for example, the case where utility (without public provision) is given by $u(c, e, y) = f(c, y) + \beta\delta h(e, y)$,where δ is a discount factor. With $\beta < 1$, the households tend to undervalue the benefits of e.

Instead of concentrating on a specific example, we follow Seade (1980) and Kanbur et al. (2006) and work with a general paternalistic social welfare function $P(x, y, g)$, which can, in principle, differ in any direction from individual utility $v(x, y, g)$.[11] Let us denote the marginal rate of substitution between g and x as $s_g = \dfrac{v_g}{v_x}$ with subscripts referring to partial derivatives. For ease of interpretation, we concentrate right from the beginning on the case where $s_g^P := \frac{P_g}{P_x} > \frac{v_g}{v_x} =: s_g$ so that from the social welfare point of view, individuals undervalue g. This is also the way Schroyen (2005) finds useful in thinking about merit goods. In the example above, the social welfare function would be one without hyperbolic discounting ($\beta = 1$), i.e. $P(c, e, y) = f(c, y) + \delta h(e, y)$.[12]

Denote the individual marginal rate of substitution between consumption and income by $s_y = \frac{v_y}{v_x}$ and the social marginal rate of substitution by $s_y^P = \frac{P_y}{P_x}$. There is no need for the two to be equal. If labour supply and publicly-provided private goods are complements,[13] $s_y^P > s_y$, and the government would like the individual to supply more labour than he or she would typically decide

[10] Of course, in a two good model, if e is chosen incorrectly, so is x because of the budget constraint. We refer to the idea that if x was a vector of private goods, the household can choose among them without mistakes.

[11] Although we will return to the example in Section 4.

[12] Here we take a straightforward solution and equate the social welfare with ex ante optimality for the individual. Note, however, that there is a debate in the literature regarding trading off utility from consumption in different periods. For this, see Gul and Pesendorfer (2001, 2007).

[13] By complementarity we mean that the demand solved from the households' first-stage optimisation problem increases with labour, $\frac{\partial g(x, \frac{y}{w})}{\partial l} \geq 0$. As shown in Blomquist and Christiansen (1998a), this condition is equivalent to $\frac{\partial s_g}{\partial l} \geq 0$, i.e. the marginal valuation of the good g increases with hours of labour supplied. Accordingly, if $\frac{\partial g(x, \frac{y}{w})}{\partial l} < 0$, we say that publicly provided good and labour are substitutes.

himself. One interpretation of this case is that education, health and the like improve the income-earning ability and desire of households. On the other hand, one could also imagine that irrationality is related to workaholism, i.e. to a tendency to overwork. Then $s_y^P < s_y$.[14]

The social planner is assumed to have a utilitarian social welfare function over P^1 and P^2. The first constraint is a resource constraint $\sum T(y^i) = rg$ implying that the tax revenue must equal the costs of public provision, rg (where r is the marginal rate of transformation between c and g). The social planner is also restricted by self-selection constraints requiring that each household should be at least as well off by choosing the bundle of income and consumption meant for them than by mimicking the choice of the other household type. We concentrate on a case where the redistribution is preferred from high-skilled households towards low-skilled households. Thus, the only binding self-selection constraint is that the high-ability type must not mimic the choice of the low-ability type: $v^2(x^2, y^2, g) \geq v^2(x^1, y^1, g)$. Note that while the social welfare depends on P, the self-selection constraint is similar to the standard model and depends on the utility function generating private behaviour, v.

The Lagrangian of the optimisation problem is given by

$$L = N^1 P^1(x^1, y^1, g)^2 + N^2 P^2(x^2, y^2, g) + \lambda \left[v^2(x^2, y^2, g) - v^2(x^1, y^1, g)\right] + \gamma \sum_{i=1}^{2} \left(N^i \left(y^i - x^i\right) - rg\right) \quad (3)$$

and the first-order conditions by

$$N^2 P_x^2 + \lambda v_x^2 - \gamma N^2 = 0 \quad (4)$$

$$N^2 P_y^2 + \lambda v_y^2 + \gamma N^2 = 0 \quad (5)$$

$$N^1 P_x^1 - \lambda \widehat{v}_x^2 - \gamma N^1 = 0 \quad (6)$$

[14] Hamermesh and Slemrod (2004) provide a detailed analysis of workaholism.

$$N^1 P_y^1 - \lambda \widehat{v}_y^2 + \gamma N^1 = 0 \tag{7}$$

where λ and γ are Lagrange multipliers for the self-selection constraint and the government budget constraint, respectively, and $\widehat{v}$ refers to the mimicker (a type 2 representative mimicking the choice of type 1).

2.1 Marginal tax rates

The individuals choose the labour supply by maximizing their utility subject to the budget constraint $x = y - T(y)$. Assuming that the tax function is differentiable, the marginal tax rate can be expressed as $MTR = T' = \frac{v_y}{v_x} + 1$. In discrete type models, the optimal tax function often includes kinks; therefore, strictly speaking, the marginal tax rate refers to the curvature of a corresponding smooth tax function around the kink. The marginal tax rates for both household types can therefore be derived by dividing (5) by (4) and (7) by (6), respectively, and they are given by

$$MTR(y^2) = \frac{P_x^2}{\gamma}\left(s_{y,2} - s_{y,2}^P\right) \tag{8}$$

and

$$MTR(y^1) = \frac{P_x^1}{\gamma}\left(s_{y,1} - s_{y,1}^P\right) + \frac{\lambda \widehat{v}_x^2}{\gamma N^1}\left(\widehat{s_y} - s_{y,1}\right), \tag{9}$$

where $\widehat{s_y} = \frac{\widehat{v}_y}{\widehat{v}_x}$ is the marginal rate of substitution between consumption and income of the mimicker.

These results give rise to the following proposition:

Proposition 1 *The marginal tax rate for the high-skilled households is negative and the marginal tax rate for the low-skilled household has an ambiguous sign.*

When the publicly-provided good is a complement to labour supply, i.e. $s_{y,i}^P > s_{y,i}$, the marginal tax rate for the high-skilled individual is negative, as the government wants to boost the labour supply and indirectly induce the individual to consume more of e. A similar effect is present in the rule for the marginal tax rate for the low-skilled (the first term at the right of (9)), but the rule also depends on a comparison between the marginal rate of substitution of a mimicker and a true type 1 household. This comparison can be signed on the basis of the single-crossing condition.[15] Therefore, $\widehat{s_y} > s_{y,1}$ and the last term in (9) is positive. The overall sign of the marginal tax rate for the low-skilled household remains ambiguous when the publicly-provided good is a complement to the labour supply.

The gist of these results is that a potential connection between the publicly-provided good and the labour supply affects the income tax rules as well. Unlike in a standard model (such as Boadway and Marchand 1995), the income tax rules depend here on the decision on public provision. This result is also derived, albeit in a different setting, by Blomquist and Christiansen (2005), where public provision (daycare) is directly proportional to labour supply and, for the effective marginal tax rate, by Blomquist and Micheletto (2006).

2.2 Optimal public provision rule

Consider next the welfare impacts of public provision. The derivative of (3) with respect to g can be written as follows:

$$\frac{dL}{dg} = N^1 P_g^1 + N^2 P_g^2 + \lambda v_g^2 - \lambda \widehat{v}_g^2 - \gamma r \tag{10}$$

Substitution from (4) and (6), and adding and subtracting $s_{g,x}^1$ and $s_{g,x}^2$ yields

$$\frac{dL}{dg} = \sum N^i s_{g,i} - r + N^1 \left(s_{g,1}^P - s_{g,1} \right) + \frac{N^2 P_x^2}{\gamma} \left(s_{g,2}^P - s_{g,2} \right) + \frac{\lambda}{\gamma} \widehat{v}_x^2 \left(s_{g,1}^P - \widehat{s}_g \right) \tag{11}$$

[15] Single-crossing condition implies that $-\frac{v_y}{v_x}$ decreases with skill level, since it is easier for the high ability individuals to transform labour for consumption. The condition is an implication of the agent monotonicity condition. The sufficient condition is that the consumption is not inferior, i.e. demand for goods increases with income.

where the mimicker's marginal rate of substitution is denoted by $\widehat{s}_g$.

Boadway and Marchand (1995, Proposition 3) show that when the publicly-provided good is provided more and the good is a complement to labour supply, the mimicker is crowded out first, i.e. he is forced to overconsume that good. This is because he has the same income as the true type 1 individual, but supplies less labour. He also supplies less labour than a true type 2 individual. When none of the true ability persons is crowded out, i.e. the publicly-provided amount does not exceed the amount individuals wish to buy themselves, individuals' maximisation implies that the marginal rate of substitution is equal to the marginal rate of transformation. This means that $\sum N^i s_{g,i} = r$. Therefore, the sign of $\frac{dL}{dg}$ at the starting point where income taxes are set optimally is determined by the rest of the terms. The terms $s^P_{g,1} - s_{g,1}$ and $s^P_{g,2} - s_{g,2}$ measure the deviations between the social planner's and the individual's marginal rate of substitution. The case in which we focus on is $s^P_{g,i} > s_{g,i}$.

The sign of the last term at the right of (11) depends on the relative valuation of g between a mimicker and social planner for type 1. As explained above, in a standard welfarist case, the comparison is between $s_{g,1}$ and $\widehat{s}_g$. There, if public provision and the labour supply are complements, the labour supply of a mimicker is smaller than that of true low-skilled person, and $s_{g,1} > \widehat{s}_g$. In our case the sign truly depends on the comparison between $s^P_{g,1}$ and $\widehat{s}_g$. When public provision and the labour supply are complements, the social planner wants the true type 1 individual to consume even more of z than he or she otherwise would. Therefore, $s^P_{g,1} > \widehat{s}_g$. The following proposition summarizes this discussion:

Proposition 2 *When individuals undervalue z from the viewpoint of social welfare, public provision of z is welfare improving.*

In policy terms, there can exist important examples of public provision that are useful tools to deal with both asymmetry of information and irrationality. Policies that increase income-earning abilities and labour force participation, but which individuals tend to undervalue, could include investments in education and health care.

This result is related to the work by Blomquist and Micheletto (2006) who consider optimal price subsidies for merit goods in their Proposition 1. They note that when a good is, in their terminology, a global merit good, it should be subsidised. This happens under same conditions when the good should be publicly provided in our case.

Finally, when the labour supply and the publicly-provided good are unrelated (preferences are separable between consumption and leisure), public provision can no longer be used as a tool to reduce incentives to mimicking. The terms related to paternalism nevertheless remain in the provision rule even under separable preferences. This can be seen more clearly when the provision rule is written as $\frac{dL}{dg} = \sum N^i s_{g,i} - 2r + (\widehat{s}_g - s_{g,1}) + \dfrac{N^i P_x^1}{\gamma}\left(s_{g,1}^P - \widehat{s}_g\right) + \dfrac{N^2 P_x^2}{\gamma}\left(s_{g,2}^P - \widehat{s}_g\right)$. When preferences are separable between consumption and leisure, the term $(\widehat{s}_g - s_{g,1})$ vanishes from the rule.

Corollary 1 *When preferences are separable between consumption and leisure and individuals undervalue z, public provision of z is welfare improving.*

3 Differences in rationality

There is, of course, no need to require that all individuals are equally irrational. As in O'Donoghue and Rabin (2003), it is more reasonable to assume that irrational individuals form only a part of the population. Given that innate rationality is something the social planner cannot directly observe, the question is how public policy should be formulated to account for both the perfectly and the less-than-perfectly rational households.

Combining this extension with the earlier assumption that households differ in their income-earning ability leaves us with four different types of households. Because of the complications of the problem, we follow Blomquist and Christiansen (2003, 2004) and concentrate on a somewhat simpler three-type model. We assume, quite plausibly, that type 1 individuals with low wage rates are all irrational in their choices over publicly-provided goods. Part of the high-income earners are assumed to be irrational (type 2) and part fully rational (type 3). For the latter class, $P = v$. An alternative plausible case would be one where some, but not all, low-income earners are irrational whereas all high-income earners are rational. We will return the consequences of this alternative after presenting the results. The number of individuals in each group is given by $N^i(> 0)$, such that $\sum N^i = 1$.

3.1 On self-selection constraints and equilibria

The formulation of self-selection constraints becomes more complicated. It depends, first, on how the distributional preferences of the government hinge on the rationality of households. Second, the ordering of the two high-ability households now depends on how the labour supply is related to the degree of rationality. Therefore, a large number of different cases can emerge. In order not to lose tractability, we concentrate on the case which seems most plausible: the social planner wants to redistribute from the rich to the poor, as in a standard income tax model, and the labour supply and the publicly-provided good, z, are complements. Then type 3 households, while consuming more of z, supply more labour than type 2 households. When rational individuals value more of the publicly-provided good and the good is a complement to labour supply, rational households will never want to mimic irrational households, as they would have to under-consume the publicly-provided good. When redistribution occurs from high-income earners to lower income earners, lower income earners are not willing to mimic high-income earners. For these reasons, only the downward self-selection constraints are binding.[16] As a result of assumptions about the individual utility functions and characteristics the single-crossing property holds also here, two potential self-selection constraints must be considered: type households 3 should not be allowed to mimic type 2 households, and type 2 households should not be allowed to mimic type 1 households.

Two possible equilibria may emerge: a separating equilibrium, where the social planner picks a separate bundle for all households (Figure 1), or a pooling equilibrium, where two or more types of households cannot be distinguished

[16] Blomquist and Christiansen (2003) also concentrate on downwards-binding self-selection constraints.

by the planner. We concentrate here on a kind of partial pooling equilibrium, where the planner cannot distinguish between the two lower income types, and there is a common bundle for type 1 and type 2 households (Figure 2), whereas a separate bundle for type 3 households. The possibility, that types 2 and 3 would have common bundle, does not seem plausible, as they have the same productivity and due to difference in rationality they never want to choose the same amount of consumption. We concentrate on the pooling equilibrium in the main text, while the separating equilibrium is dealt with in an Appendix.

In a pooling equilibrium, type 1 and type 2 households receive the same gross and net income. Thus, there is only one binding self-selection constraint: type 3 households should not mimic the common choice of types 1 and 2. The Lagrangian is now given by

$$\begin{aligned} L = N^1 P^1(x^1, y^1, g) + N^2 P^2(x^2, y^2, g) + N^3 v^3(x^3, y^3, g) \\ + \lambda \left[v^3(x^3, y^3, g) - v^3(x^1, y^1, g) \right] \\ + \gamma \left[\sum_h N^i \left(y^h - x^h \right) - rg \right] \end{aligned} \quad (12)$$

Remembering that $x^1 = x^2$ and $y^1 = y^2$ the first-order conditions with respect to $x^h$and y^h, $h = 1, 3$ are

$$\left(N^3 + \lambda \right) v_x^3 - \gamma N^3 = 0 \quad (13)$$

$$\left(N^3 + \lambda \right) v_y^3 + \gamma N^3 = 0 \quad (14)$$

$$N^1 P_x^1 + N^2 P_x^2 - \lambda \widehat{v}_x^3 - \gamma (N^1 + N^2) = 0 \quad (15)$$

$$N^1 P_y^1 + N^2 P_y^2 - \lambda \widehat{v}_y^3 + \gamma (N^1 + N^2) = 0 \quad (16)$$

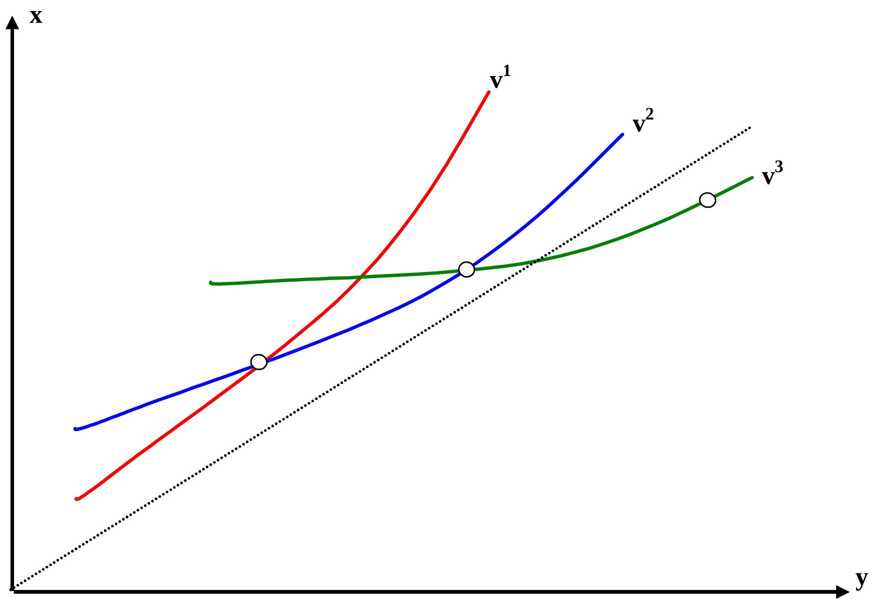

Figure 1: Separating equilibrium

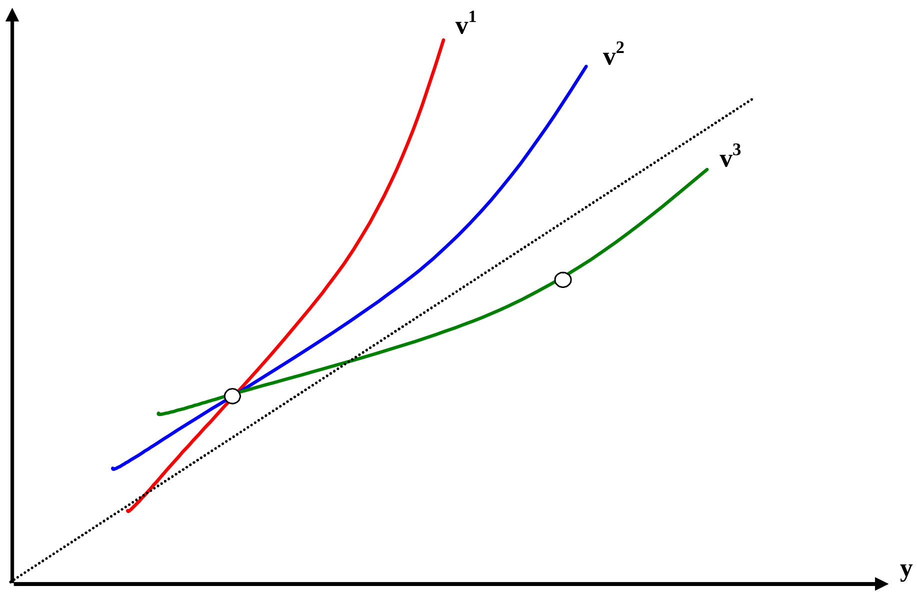

Figure 2: Pooling equilibrium

3.2 Marginal tax rates

It is straightforward to see from the first-order conditions (13) and (14) that the marginal tax rate of type 3 is zero. Marginal tax rates for the other types can be derived by adding and subtracting terms v_x^1 and v_y^1 into Equations (15) and (16), respectively, and dividing (16) by (15). The marginal tax rates of types 1 and 2 can be written as

$$\begin{aligned} MTR^1 = MTR^2 = & \frac{\lambda \widehat{v}_x^3}{\gamma(N^1+N^2)}(\widehat{s}_{y,3} - s_{y,1}) \\ & + \frac{N^1 P_x^1}{\gamma(N^1+N^2)}(s_{y,1} - s_{y,1}^P) + \frac{N^2 P_x^2}{\gamma(N^1+N^2)}(s_{y,2} - s_{y,2}^P) \quad (17) \end{aligned}$$

There are two channels present in the income tax rate of the irrational types: the first term arises from the self-selection constraint and the last two terms from irrationality. The first term is positive, as $\widehat{v}_x^3 > 0$ and mimicker's marginal rate of substitution is greater than a representative of true type 1 as a result of the single-crossing property. The last two terms depend on the difference between the true type's and the social planner's marginal rates of substitution. When the publicly-provided good and the labour supply are complements, $s_y^P > s_y$. It is also natural to assume that even in a three-type case, both P_x^1 and P_x^2 are positive. The rule for marginal tax rates in the pooling optimum can thus be stated in the following way

Proposition 3 *The marginal tax rate of the rational high-productivity type is zero and the sign of the marginal tax rate of the other two types is ambiguous.*

Note finally that in a special case where $P_x^1 = P_x^2$, i.e. the government's valuation depends only on income level, not rationality, the rule is the same as in the two-type case.[17]

[17] This can be seen by combining (15) and (16) to obtain (9).

3.3 Public provision

To find the optimal rule for public provision, we derive the derivative of Lagrangian (12) with respect to g.

$$\frac{dL}{dg} = N^1 P_g^1 + N^2 P_g^2 + \left(N^3 + \lambda\right) v_g^3 - \lambda \widehat{v}_g^3 - \gamma r \tag{18}$$

After the same kind of manipulation as in the previous section, this can be written as

$$\begin{aligned} \frac{dL}{dg} = \sum_i N^i s_{g,i} - r \\ + N^1 \left(s_{g,1}^P - s_{g,1}\right) + N^2 \left(s_{g,2}^P - s_{g,2}\right) \\ + N^1 \left[\frac{P_x^1}{\gamma} - 1\right] \left(s_{g,1}^P - \widehat{s}_{g,3}\right) + N^2 \left[\frac{P_x^2}{\gamma} - 1\right] \left(s_{g,2}^P - \widehat{s}_{g,3}\right) \end{aligned} \tag{19}$$

Consider first which type will be crowded out when public provision is increased. Unlike in the two-type case, it is not clear that the mimicker is the first to become crowded out. The mimicker enjoys more leisure than a true type-1 individual, and when the publicly-provided good is a complement to labour supply, he tends to suffer from an increase in the provision of that good. On the other hand, the mimicker is fully rational, and therefore, he inherently values the publicly-provided good more than the irrational type 1. In comparison to type 2, he and the mimicker have the same amount of leisure, but the mimicker is rational. That is why a true type 2 individual becomes crowded out earlier than the mimicker. This implies at least one of the true types is forced to overconsume the publicly-provided good, if its provision is increased, and hence, the first row of Equation (19) becomes negative.

In the second row, there are two terms arising from irrationality. As the social planner values the consumption of the publicly-provided good more than types 1 and 2, these terms are positive.

The terms in the last row arise from mimicking. Consider first the multipliers $\frac{P_x^h}{\gamma} - 1$. Following Blomquist and Christiansen (2004), very useful expressions can be obtained by differentiating the Lagrange function in (12) with respect to x^1 and x^2: $\frac{\partial L}{\partial x^h} = \frac{P_x^h}{\gamma} - 1$, $h = 1, 2$. These expressions represent the welfare effect of a hypothetical increase in net income x^h of household h. If the desired direction of redistribution is from an irrational high ability type towards an irrational low ability type, it can be concluded that $\frac{\partial L}{\partial x^1} > 0$, and $\frac{\partial L}{\partial x^2} < 0$.

To deduce the sign of the terms, we concentrate on a case where g is complement to labour supply. The last term in (19) includes a comparison between $s_{g,2}^P$ and $\widehat{s}_{g,3}$. In general, this comparison cannot be signed. However, note that both the mimicker and the true type 2 representative are high-skilled, and therefore, they supply an equal amount of labour. If the government corrects the valuation of type 2 exactly to the fully rational level, $s_{g,2}^P$ is as large as $\widehat{s}_{g,3}$, and the last term vanishes from the rule.

In the first of the terms in the last row of (19), the sign of the comparison between $s_{g,1}^P$ and $\widehat{s}_{g,3}$ is ambiguous. The true type 1 representative works more, and, since g and the labour supply are complements, tends to favour g more. On the other hand, the type 3 individual is rational and therefore also favours g. As opposed to the case in the previous paragraph, since the two individuals have different skill levels, it is not clear that the government wants to impose the same marginal rate substitution on them.

Based on the discussion of all the terms in Equation (19) , one can deduce the following proposition:

Proposition 4 *In a pooling equilibrium, even when consumption of good g is too low from the social welfare point of view, public provision of g is not unambiguously welfare improving.*

When the publicly-provided private good and the labour supply are complements, both the aim to reduce the harm from irrationality and asymmetric information separately would speak for a positive level of public provision. However, when these two aims are put together, it is no longer clear that public provision improves welfare both as a tool to correct irrationality and asymmetric information. Taking into account differences in rationality, the two problems, irrationality and asymmetry of information, interact in a way that precludes stating simple policy rules without further assumptions. Note, however, that a full analysis of this would require simulations to obtain knowledge about the likely relative magnitudes of the different terms in the provision rule. Therefore, the desirability of public provision does not necessarily need to be ambiguous.

Jordahl and Micheletto (2005) have derived results that have a similar character. They consider a model with taste differences and differences in income-earning abilities and show that, in a three-type model, the commodity tax can have an ambiguous sign even if the sign was unambiguous without taste differences.

Consider finally the case when preferences are separable between the labour supply and the publicly-provided private good. In pooling equilibrium we have $s_{g,1} = s_{g,2} =: s_g^*$. In that case rule (19) reduces to $\frac{dL}{dg} = \sum_i N^i s_{g,i} - r + \frac{N^i P_x^1}{\gamma}\left(s_{g,1}^P - s_g^*\right) + \frac{N^2 P_x^2}{\gamma}\left(s_{g,2}^P - s_g^*\right) + \left(N^1 + N^2\right)\left(\widehat{s}_{g,3} - s_g^*\right)$. The two positive terms in the middle would vanish only when the agents are rational, i.e. $s_g^P = s_g^*$. The last term is positive, as with separable preferences rational type 3 mimicker values public provision of z more than irrational types. But notice, that assuming separable preferences do not change the fact that a true type 2 individual – having a smaller taste for the publicly-provided good – is the first to become crowded out. Therefore, $\sum_i N^i s_{g,i} - r < 0$, and one can note the following:

Corollary 2 *Even when preferences are separable between consumption and labour supply, public provision of good g is not necessarily welfare improving.*

The question arises how sensitive these results are to the assumption of pooling equilibrium and to our assumption of how income-earning ability and rationality are related. The Appendix presents results for a separating equilibrium, where each households faces a distinct tax/benefit package, with the same types as above. The main point that public provision is not unambiguously welfare improving with differences in rationality remains valid in the separating case as well. Turning to an alternative case, where there are rational and irrational low-income earners, but all high-income earners are rational, it appears that public provision is not unambiguously welfare-improving, similarly to the result above.[18]

Solving whether a separating or some of pooling equilibria is optimal is a complicated task, probably requiring numerical computations, and is left for further work. The solution could depend on the relative shares of households of different types. If, for example, the share of type 2 persons is negligible, pooling type 1 and type 2 households into a same group is likely to be useful.[19]

Finally, as earlier literature has pointed out (starting from Edwards et al., 1994), price subsidies can be used as an alternative mechanism to public provision to reach the same goals. In a longer discussion paper version (Pirttilä and Tenhunen 2005), we consider the optimal commodity tax (subsidy) rule in two cases: the benchmark case of two types and the pooling equilibrium in the three-type case. It turned out that the results are analogous to public provision: introducing price subsidies is not unambiguously welfare improving, but it depends on similar types of terms as the optimality of public provision considered here.

4 An example: Health behaviour

The frequency of habits or conditions commonly known to be bad for our health, e.g. obesity, smoking or the use of narcotic drugs, proves that people make decisions that turn out to be bad for their health in the future. According to OECD health statistics, 15-30% of adults in OECD smoke cigarettes daily and adult obesity in the US has increased from 15% in 1980 to 30% in 2003. Needless to say, these trends have a high direct cost on the public sector – Finkelstein, Fiebelkorn and Wang (2003) estimate that the costs due to being overweight and obesity accounted for 9% of total medical expenses in the US at the end of 1990s – but it is also clear that bad health has indirect effects

[18] The optimality of public provision with an alternative composition of the economy is considered in detail in an additional Appendix available from the authors upon a request. The main result there is that in both separating and pooling equilibrium, public provision of g is not unambiguously welfare improving when the model consists of rational and irrational low-income earners and rational high-income earners.

[19] We are grateful to a referee for pointing this out.

through reduced labour supply and tax revenues.

There are a number of reasons why these outcomes may, at least, partly result from mistakes in human behaviour. First, disadvantageous choices like smoking or eating unhealthy food cause health problems mainly in the future. People may undervalue the future costs due to 'distortions' in discounting, as hyperbolic discounting models in both psychology (Ainslie and Haslam, 1992) and economics (Laibson, 1997) demonstrate.

Second, decisions affecting health can be subject to projection bias, according to which people mispredict their future preferences when comparing them to the current preferences (Loewenstein, O'Donoghue and Rabin, 2003). For example, people tend to exaggerate the need for food when they go shopping when hungry or undervalue the adaptation to changes in life circumstances, such as medical conditions.[20] Projection bias may also partly explain dynamic inconsistency, like evidence from health clubs shows: People choose disadvantageous contracts with the health clubs as they mispredict their preference to visit them. With agents making mistakes in discounting, this can be interpreted as a trial to commit themselves to increase future attendance to promote health (DellaVigna and Malmendier 2004, 2006).[21]

We illustrate one of these linkages, namely short-sightedness, in the context of our model. Consider therefore our earlier example of Section 2, where individual preferences between the private good, c, and health, e, are given by $u(c, e, y) = f(c, y) + \beta\delta h(e, y)$, where δ is a discount factor and $\beta < 1$ (hyperbolic preferences). The social planner does not suffer from myopia and its preferences are otherwise the same, but with $\beta = 1 : f(c, y) + \delta h(e, y)$. From individual optimisation, $v_x = u_c = f_c$ and $v_g = u_e = \beta\delta h_e$. Similarly, $P_x = u_c$ and $P_g = \delta h_e$. Consider now the rule for optimal public provision with no differences in rationality, given in equation (11). With the preferences above, it can be rewritten as

[20] For a summary of evidence on projection bias, see Loewenstein and Schkade (1999).

[21] Non-standard behaviour can also be explained by social attitudes, medical addictions or illnesses, or the sheer knowledge requirement may also be too severe. (See e.g. Ajzen and Fishbein (1980) for a theory of reasoned action, and Rice (2002) and Richman (2005) for discussion from the perspective of health economics.)

$$\frac{dL}{dg} = \sum_{i=1,2} \frac{N^i \beta \delta h_e^i}{f_c^i} - r + \sum_{i=1,2} (1-\beta) N^i \frac{\delta h_e^i}{f_c^i} + \frac{\lambda}{\gamma} \widehat{f}_c^2 \left(\frac{\delta h_e^1}{f_c^1} - \frac{\beta \delta \widehat{h}_e^2}{\widehat{f}_c^2} \right) \tag{20}$$

Here, the terms that reflect the difference between the social planner's and individual's preferences only differ with respect to the discount factor. Since $\beta < 1$, the public provision rule corrects the individuals' mistake in investing in health, due to myopia. The terms referring to mimicking remain in the rule, as we have assumed no separability in the formulation of the utility function. Since investments in health are likely to be complementary with labour supply, the last term is positive, and both it and the term $(1-\beta)\frac{\delta h_e^i}{f_c^i}$ work towards optimally positive public provision.

Similarly, in the presence of differences in rationality, the public provision rule can be written as

$$\frac{dL}{dg} = \sum_{i=1,2} \frac{N^i \beta \delta h_e^i}{f_c^i} + \frac{N^3 \delta h_e^3}{f_c^3} - r + \sum_{i=1,2} (1-\beta) N^i \frac{\delta h_e^i}{f_c^i} + \left[\frac{f_c^1}{\gamma} - 1 \right] \left(\frac{\delta h_e^1}{f_c^1} - \frac{\delta \widehat{h}_e^3}{\widehat{f}_c^3} \right) \tag{21}$$

The last term in (19) vanishes since the government corrects the choice for type 2 to correspond exactly to the fully-rational choice of type 3. While the term $(1-\beta)\frac{\delta h_e^i}{f_c^i}$ increase the benefits of public provision, the sign of the last term in (21) is ambiguous: the valuation of health investment by a low-skilled individual 1, even corrected by the social planner, is not necessarily equal to the valuation of a highly-skilled, fully-rational, type 3 individual.

5 Conclusions

Much of modern governments' activities are not related to night-watchman duties or the provision of pure public goods. A major share of public expenditure is directed to the funding of publicly-provided private goods, such as education, health care, care of the elderly, and pension policies. Using a modern, information-based, optimal tax framework, this paper considered two motives for the public provision of private goods: redistribution and paternalism. The latter concern is warranted by recent discussion on behavioural economics indicating that individuals may have a tendency towards biased decisions related to these goods. This was illustrated in the paper by an application to health decisions.

Informational asymmetry between the government and the taxpayers implies that incentive effects must be taken into account in redistributive tax policy and the design of public services. Irrational behaviour by individuals provides a potential scope for a paternalistic policy where the government, through its tax and public provision policy, induces behaviour that is closer to what it believes is truly best for the individuals. Our results show that public provision can, in principle, be used as a mechanism to address these concerns. This is the case, in a model where all households suffer from similar bounded rationality, if the social planner wants to induce individuals consume more publicly-provided goods than they themselves would buy and if the publicly-provided private good is a complement to the labour supply. Then, the paternalistic considerations and a desire to alleviate distortions from income taxation by public provision are aligned.

However, in a richer model where some of the households are fully rational, redistributive and paternalistic objectives can clash even if public provision boosts the labour supply. This result implies that the two different aims can interact in a way that precludes making simple policy conclusions. Part of the intuition for this unexpected result stems from the idea that the government may want to value the utility of irrational households less than that of fully-rational households. On the other hand, if redistribution is directed from rich to poor and the fully-rational households work more, the government would like to favour the poorer irrational households. And finally, if mimickers are fully rational while those who are mimicked are irrational, mimickers suffer less from overprovision of public goods. This, in turn, is not desirable for redistribution. Therefore, the two policy objectives point to different directions.

The decision of a public good provision also affects the rules governing the marginal tax rates. This is the case if the labour supply and the demand for publicly-provided private good are related. Then the optimal marginal tax rate rules include novel corrective terms.

Subject to the qualifications presented above, it is not entirely implausible that a case exists for public provision of certain goods for paternalistic reasons, but paternalistic aims can interact with other policy objectives producing somewhat surprising results. The gain from paternalism must be weighted against the potential mistakes the social planner can make when designing paternalistic policies, difficulties in welfare economics with time-dependent preferences, and the potential for the market to innovate new self-control mechanisms. It is also important to bear in mind the ethical dimensions involved. A paternalistic policy that, e.g. protects the individuals is welcomed by some, while others cannot accept such restrictions on individual freedom.

References

Ainslie, G. and N. Haslam (1992) ‘Hyperbolic Discounting’ in Loewenstein G. and J. Elster (Eds.) Choice over Time. Russel Sage Foundation Press, New York.

Ajzen, I. and M. Fishbein (1980) ‘Understanding Attitudes and Predicting Social Behaviour’. Prentice Hall.

Aronsson, T. and L. Thurström (2005) ‘Optimal Paternalism: Sin Taxes and Health Subsidies’, mimeo, University of Umeå.

Besley, T. (1988) ’A Simple Model for Merit Good Arguments’, Journal of Public economics 35, 371-384.

Blomquist, S. and V. Christiansen (1998a) ’Topping Up or Opting Out? The Optimal Design of Public Provision Schemes’, International Economic Review 39, 399-411.

Blomquist, S. and V. Christiansen (1998b) ‘Price Subsidies versus Public Provision’, International Tax and Public Finance 5, 283-306.

Blomquist, S. and V. Christiansen (2003) ‘Is there a Case for Public Provision of Private Goods if Preferences are Heterogeneous? An Example with Day Care’, CESifo Working Paper No. 938.

Blomquist, S. and V. Christiansen (2004) ‘Taxation and Heterogeneous Preferences’, CESifo Working Paper No. 1244.

Blomquist, S. and V. Christiansen (2005) ‘Welfare Enhancing Marginal Tax Rates: The Case of Publicly Provided Day Care’, Uppsala University, mimeo.

Blomquist, S. and L. Micheletto (2006) ‘Optimal Redistributive Taxation When Government’s and Agents’ Preferences Differ’, Journal of Public Economics 90, 1215-1233.

Boadway, R. and M. Keen (1993) ‘Public Goods, Self-selection and Optimal Income Taxation’, International Economic Review 34, 463-478.

Boadway, R. and M. Marchand (1995) ‘The Use of Public Expenditure for Distributive Purposes’. Oxford Economic Papers 47, 45-59.

Boadway, R., M. Marchand, P. Pestieau and M. del mar Racionero (2002) ‘Optimal Redistribution with Heterogeneous Preferences for Leisure’, Journal of Public Economic Theory 4, 475-498.

Camerer, C.F. and G. Loewenstein (2004) ‘Behavioral Economics: Past, Present, Future’ in C.F. Camerer, G. Loewenstein and M. Rabin (eds.) Advances in Behavioral Economics. Princeton: University Press.

Cremer, H. and Gahvari, F. (1997) ‘In-kind Transfers, Self-selection and Optimal Tax Policy’. European Economic Review 41, 97-114.

DellaVigna, S. and U. Malmendier (2004) ‘Contract Design and Self-Control: Theory and Evidence’. Quarterly Journal of Economics 119, 353-402.

DellaVigna, S. and U. Malmendier (2006) ‘Paying Not To Go To the Gym’. American Economic Review 96, 694-719.

Edwards, J., Keen, M. and Tuomala, M. (1994) ‘Income tax, Commodity Taxes and Public Good Provision’, FinanzArchiv 51, 472-487.

Finkelstein, E. A., I. C. Fiebelkorn and G. Wang (2003) ‘National Medical Spending Attributable to Overweight and Obesity: How Much and Who’s Paying?’ Health Affairs W3, 219-226.

Glaeser, E.L. (2005) ’Paternalism and Psychology’, NBER Working Paper No. 11789.

Gruber, J. and B. Köszegi (2004) ‘Tax Incidence when Individuals are Time-inconcistent: The Case of Cigarette Excise Taxes’, Journal of Public Economics 88, 1959-1987.

Gul, F. and W. Pesendorfer (2001) ’Temptation and Self-Control’, Econometrica 69, 1403-1435.

Gul, F. and W. Pesendorfer (2007) ’Harmful Addiction’, Review of Economic Studies 74, 147-172.

Hamermesh, D. and J. Slemrod (2004) ‘The Economics of Workaholism: We Shouldn’t Have Worked on This Paper’, University of Michigan, Mimeo.

Jordahl, H. and L. Micheletto (2005) ‘Optimal Utilitarian Taxation and Horizontal Equity’, Journal of Public Economic Theory 7,681-708.

Kanbur, R., J. Pirttilä and M. Tuomala (2006) ‘Non-Welfarist Optimal Taxation and Behavioral Public Economics’, Journal of Economic Surveys 20, 846-868.

Köszegi, B. (2005) ’On the Feasibility of Market Solutions to Self-Control Problems’, Swedish Economic Policy Review 12, 71-94.

Laibson, D. (1997) ‘Golden Eggs and Hyperbolic Discounting’, Quarterly Journal of Economics 116, 443-477.

Loewenstein, G., T. O’Donoghue and M. Rabin (2003) ’Projection Bias In Predicting Future Utility’, The Quarterly Journal of Economics, 118(4), 1209-1248

Loewenstein, G. and D. Schkade (1999) ’Wouldn’t it be nice? Predicting Future Feelings’ in Kahneman, D., E. Diener and N. Schwartz (Eds.) Well-being: The Foundations of Hedonic Psychology. Russel Sage Foundation Press, New York.

Le Grand, J. (2003) Motivation, Agency, and Public Policy. Of Knights & Knaves, Pawns & Queens, Oxford: University Press.

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

Musgrave, R. (1959) ’The Theory of Public Finance’, New York: McGraw-Hill Book Company,

O’Donoghue, T. and M. Rabin (2003) ‘Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes’, American Economic Review 93, 186-191.

Pirttilä, J. and S. Tenhunen (2005) 'Pawns and Queens Revisited: Public Provision of Private Goods when Individuals Make Mistakes', CESifo Working Paper No. 1466.

Pirttilä, J. and Tuomala, M. (2002) 'Publicly Provided Private Goods and Redistribution: A General Equilibrium Analysis', Scandinavian Journal of Economics 104,173-188.

Racionero, M. del Mar (2001) 'Optimal Tax Mix with Merit Goods', Oxford Economic Papers 53, 628-641.

Rice, T. (2002) 'The Economics of Health Reconsidered, 2nd edition', Health Admimistration Press, Chigago.

Richman, B. D. (2005) 'Behavioral Economics and Health Policy: Understanding Medicaid's Failure'. Cornell Law Revew 90, 3, 705-768.

Sandmo, A. (1983) 'Ex Post Welfare Economics and the Theory of Merit Goods', Economica 50, 19-33.

Schroyen, F. (2005) 'An Alternative Way to Model Merit Good Arguments', Journal of Public Economics 89, 957-966.

Seade, J. (1980) 'Optimal Non-linear Policies for Non-utilitarian Motives', in D. Collard, R. Lecomber and M. Slater (eds.) Income Distribution: The Limits to Redistribution, Bristol: Scientechnica.

Sheshinski, E. (2003) 'Optimal Policy to Influence Individual Choice Probabilities', The Hebrew University of Jerusalem, mimeo.

Stern, N. (1982) 'Optimal 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.

Tobin, J. (1970) 'On Limiting the Domain of Inequality', Journal of Law and Economics 13, 263-277.

Appendix: Separating equilibrium

The other possible case is a separating equilibrium where each household is given a separate bundle of x and y. Now there are two binding self-selection constraints: type 3 households should prefer his bundle to type 2 household's choice and type 2 households should not want to mimic the choice of type 1 households. The resulting Lagrange function is

$$L = N^1 P^1(x^1, y^1, g) + N^2 P^2(x^2, y^2, g) + N^3 v^3(x^3, y^3, g) + \lambda \left[v^3(x^3, y^3, g) - v^3(x^2, y^2, g)\right] + \delta \left[v^2(x^2, y^2, g) - v^2(x^1, y^1, g)\right] + \gamma \left[\sum_h N^i \left(y^h - x^h\right) - rg\right] \quad (22)$$

and the first-order conditions are given by

$$N^3 v^3_x + \lambda v^3_x - \gamma N^3 = 0 \quad (23)$$

$$N^3 v^3_y + \lambda v^3_y + \gamma N^3 = 0 \quad (24)$$

$$N^2 P^2_x - \lambda \widehat{v}^3_x + \delta v^2_x - \gamma N^2 = 0 \quad (25)$$

$$N^2 P^2_y - \lambda \widehat{v}^3_y + \delta v^2_y + \gamma N^2 = 0 \quad (26)$$

$$N^1 P^1_x - \delta \widehat{v}^2_x - \gamma N^1 = 0 \quad (27)$$

$$N^1 P^1_y - \delta \widehat{v}^2_y + \gamma N^1 = 0 \quad (28)$$

By standard procedures, the following marginal tax rates can be derived from first-order conditions

$$MTR(y^3) = 0 \quad (29)$$

$$MTR(y^2) = \frac{P_x^2}{\gamma}\left(s_2 - s_2^P\right) + \frac{\lambda \widehat{v}_x^3}{\gamma N^2}\left(\widehat{s}_3 - s_2\right), \tag{30}$$

$$MTR(y^1) = \frac{P_x^1}{\gamma}\left(s_1 - s_1^P\right) + \frac{\lambda \widehat{v}_x^2}{\gamma N^1}\left(\widehat{s}_2 - s_1\right), \tag{31}$$

On the basis of equations (29), (30) and (31), one notes that if the publicly-provided good is a complement to the labour supply (i.e. $s < s^P$), then the marginal tax rate for the type 3 households is zero, and the marginal tax rates for the type 1 and type 2 households have ambiguous signs. Note first that since we assume type 3 households are rational, there is no need to distort their choice. The marginal tax rates for the type 1 and 2 households consist of two terms, a corrective term and a standard self-selection term. As in the two-type model, these have opposite signs, and the overall signs of the marginal tax rates for 1 and 2 remain ambiguous.

Consider next the welfare impacts of public provision. The derivative of (22) with respect to g can be written as follows:

$$\frac{dL}{dg} = N^1 P_g^1 + N^2 P_g^2 + N^3 v_g^3 + \lambda v_g^3 - \lambda \widehat{v}_g^3 + \delta v_g^2 - \delta \widehat{v}_g^2 - \gamma r \tag{32}$$

Following a similar procedure as in earlier sections, equation (32) can be written as $s_{g,2}$

$$\begin{aligned} \frac{dL}{dg} = \sum N^i s_{g,i} - r \\ + N^1\left(s_{g,1}^P - s_{g,1}\right) + \left[N^2 - \frac{\delta v_x^2}{\gamma}\right]\left(s_{g,2}^P - s_{g,2}\right) \\ + \frac{\delta \widehat{v}_x^2}{\gamma}\left(s_{g,1}^P - \widehat{s}_{g,2}\right) + \frac{\lambda \widehat{v}_x^3}{\gamma}\left(s_{g,2}^P - \widehat{s}_{g,3}\right) \end{aligned} \tag{33}$$

The interpretation of (33) follows the ones in Sections 2 and 3. Start by considering the first row. Now there are two mimickers. In comparison to the type 3 mimicker, the true type 2 individual becomes crowded out earlier as they have same amount of leisure but type 3 mimicker, being fully rational, values z more. Type 2 mimicker and true type 1 have the same rationality, but type 2 mimicker enjoys more leisure. Thus, when publicly-provided good is complement with labour supply, type 2 mimicker becomes crowded out before true type 1 and the first row is negative.

The differences $s_{g,h}^{P} - s_{g,h}$, $h = 1, 2$, are positive. $s_{g,1}^{P}$ is likely to be higher than $\widehat{s}_{g,2}$. Both 1 and 2 are irrational and there are therefore no taste differences, whereas the type 1 individual works more. In general, $s_{g,2}^{P}$ can be smaller or larger than $\widehat{s}_{g,3}$. Following the reasoning in the main text, they are, however, equal if the social planner corrects the valuation of type 2 exactly to the level chosen by the rational type 3 mimicker. In this case, the last term vanishes from the rule.

The sign of the coefficient $N^2 - \frac{\delta v_x^2}{\gamma}$ is not trivial. Using the first-order condition (25) we find that $N^2 - \frac{\delta v_x^2}{\gamma} = \frac{N^2 P_x^2}{\gamma} - \frac{\lambda v_x^3}{\gamma}$. This is negative when $\frac{N^2 P_x^2}{\gamma} < \frac{\lambda v_x^3}{\gamma}$. This condition holds when the social planner's valuation of the income of type 2 is sufficiently small. Overall, all terms in (33) are positive or zero, except $\left[N^2 - \dfrac{\delta v_x^2}{\gamma}\right] \left(s_{g,2}^{P} - \widehat{s}_{g,3}\right)$. According to this discussion, even when consumption of good g is too low from the social welfare point of view and g is complement to the labour supply, public provision of g is not unambiguously welfare improving in a separating equilibrium.

The irrationality of type 1 household, implied by the term $s_{g,1}^{P} - s_{g,1}$ leads to an unambiguously positive public provision and, thus, the higher the distortion due to irrationality is, the higher the public provision level should be. However, when the term referring to the irrationality of type 2 is negative, it implies that the higher the gap between the socially desirable and the actual level of consumption of type 2 is, the lower the public provision level should be. This can be interpreted so that if government does not value the consumption of type 2 agents enough (i.e. if $\frac{N^2 P_x^2}{\gamma} < \frac{\lambda v_x^3}{\gamma}$), their irrationality problem is worsened by public provision. As in the pooling equilibrium, there is again a potential conflict between the two policy goals of the government (correction of asymmetric information and irrationality).

Correction to Chapter 4

There are few unfortunate misprints and errors in the text. These mistakes do not change any other results or interpretations in the paper.

- In Equation (3) there is an excess superscript in the first term. It should be

$$L = N^1 P^1(x^1, y^1, g) + N^2 P^2(x^2, y^2, g) + \lambda \left[v^2(x^2, y^2, g) - v^2(x^1, y^1, g) \right] + \gamma \sum_{i=1}^{2} \left(N^i \left(y^i - x^i \right) - rg \right)$$

- On page NN, int he paragraph before Corollary 1 there is an incorrect superscript in the formula. It should be

$$\frac{dL}{dg} = \sum N^i s_{g,i} - 2r + (\widehat{s}_g - s_{g,1}) + \frac{N^1 P_x^1}{\gamma} \left(s_{g,1}^P - \widehat{s}_g \right) + \frac{N^2 P_x^2}{\gamma} \left(s_{g,2}^P - \widehat{s}_g \right)$$

- In Equation (20) there should be $N^1 (1-\beta) \frac{\delta h_e^1}{f_c^1} + \frac{N^2 f_c^2}{\gamma} (1-\beta) \frac{\delta h_e^2}{f_c^2}$ instead of $\sum_{i=1,2} (1-\beta) N^i \frac{\delta h_e^i}{f_c^i}$, i.e. the correct version is

$$\frac{dL}{dg} = \sum N^i \frac{\beta \delta h_e^i}{f_c^i} - r + N^1 (1-\beta) \frac{\delta h_e^1}{f_c^1} + \frac{N^2 f_c^2}{\gamma} (1-\beta) \frac{\delta h_e^2}{f_c^2} + \frac{\lambda}{\gamma} \widehat{f}_c^2 \left(\frac{\delta h_e^1}{f_c^1} - \frac{\beta \delta \widehat{h}_e^2}{\widehat{f}_c^2} \right)$$

- In Equation (21) the coefficient N^1 in front of the last term has dropped off, i.e. it should be given as

$$\frac{dL}{dg} = \sum_{i=1,2} \frac{N^i \beta \delta h_e^i}{f_c^i} + \frac{N^3 \delta h_e^3}{f_c^3} - r + \sum_{i=1,2} N^i (1-\beta) \frac{\delta h_e^i}{f_c^i} + N^1 \left[\frac{f_c^1}{\gamma} - 1 \right] \left(\frac{\delta h_e^1}{f_c^1} - \frac{\delta \widehat{h}_e^3}{\widehat{f}_c^3} \right)$$

Chapter 5

On Optimal Lifetime Redistribution Policy

Sanna Tenhunen

University of Tampere and FDPE

and

Matti Tuomala

University of Tampere

Abstract: In this paper we examine various aspects of the optimal lifetime redistribution policy within a cohort. We characterise the optimal tax policy when society consists of individuals who do not differ only in productivity, but also in time preference or myopia. We extend Diamond's (2003) analysis on nonlinear taxation of savings into the three-type and four-type models. To gain a better understanding of the lifetime redistribution, the problem is also solved numerically.

Our results provide a rationale for distortions (upward and downward) in savings behaviour in a simple two-period model where high-skilled and low-skilled individuals have different non-observable time preferences beyond their earning capacity. If we interpret our model so that instead of private savings there is public provision of pension in the second period, then in the three-type model we find a non-monotonic pattern of the replacement rates. The numerical results suggest that retirement consumption is less dispersed than the first period consumption in a paternalistic case. The existence of myopic individuals reinforces the picture. Paternalistic government policy also increases second period consumption compared to the welfarist case.

Keywords: Optimal taxation, lifetime redistribution, heterogeneous time preferences, myopia

JEL: H21, H55.

1. Introduction

Publicly provided retirement programmes, the largest single income source of the elderly, can be justified on several grounds. Standard justifications for public pension systems are market failures generated by asymmetric information, and redistributive grounds achieved by intervening in retirement and saving decisions. A third standard justification for public pension systems or compulsory pension contributions relies on the assumption that some individuals behave myopically, consuming “excessively” during their earning years and then find themselves with insufficient savings in retirement. The first two rationales for a public pension programme have received much more attention in public economics than the third one.

There is, however, some literature on pension policies that attempts to take into account possible “undersaving” by households. Diamond (1977) discussed the case where individuals may undersave due to mistakes. Sheshinski (2003) proposed a general model with faulty individual decision-making, where restricting individuals' choices leads to welfare improvements. Feldstein (1985) studied the optimal pay-as-you-go system in an OLG setting in a case where individuals have higher discount rates than the government. Feldstein’s model was extended by İmrohoroğlu, İmrohoroğlu and Joines (2003). They made numerical simulations in a pay-as-you-go model when individuals have hyperbolic discounting preferences. They conclude that a pension system provides additional welfare for myopic individuals. Diamond and Köszegi (2003) also employ this kind of model with hyperbolic discounting to study the policy effects of endogenous retirement choices using a public pension system as a commitment device.

The third justification for a public pension system has become more motivating as the recent research in behavioural economics has demonstrated that an individual decision-making often suffers from various biases. In these situations when there is a possible conflict between a individual's long term preference and his short term behaviour the government may want to intervene. The normative analysis of such individual decision failures in the context of the design of a pension system has not yet received much attention. One important exception is the recent book by Diamond (2003). Another study by Cremer et al. (2006), closely related to our paper, also examines a non-linear social security scheme when the government has a paternalistic view and wants to help to overcome individuals' myopia problems.

As in Diamond (2003) we also analyse the lifetime redistribution of income across individuals within a cohort.[1] Thereby we avoid the dynamic complications that arise from intergenerational redistribution. More specifically, we consider a two-period variant of the Mirrlees (1971) income tax problem, where all individuals work in the first period and then retire. Mirrlees suggests that one of the limitations of his analysis is that it does not address intertemporal problems, even if "in an optimum system, one would no doubt wish to relate tax payments to the whole life pattern of income…". The Mirrlees model as an annual tax system requires that the government can commit to ignore the information that has been revealed by individual's choices. Therefore, reinterpreting the Mirrlees model as a model of lifetime redistribution, we assume that the government can commit to a lifetime tax[2]. The lifetime version of the Mirrlees model can be interpreted either

[1] Vickrey (1939) is the famous early contribution to the normative theory of lifetime income taxation. He proposed a cumulative averaging system for personal income taxation. Basing his argument on horizontal equity, Vickrey argued that if the tax schedule is convex then individuals with fluctuating incomes would pay more taxes on average than individuals with steadier incomes.

[2] In fact Roberts (1984, proposition 4) in an optimal dynamic taxation models without government commitment raises doubts about the government's ability to use information from early periods of life to accomplish redistribution in later periods of life with lower welfare cost. Gaube (2007) in turn interprets the practice of taxing current income in each period as a partial

such that the government controls first and second period consumption and labour supply directly, subject to the self-selection constraints or, alternatively, if we assume that there is no private savings we have a model of labour income taxation in the first period and public provision of pension in the second period. The first interpretation means that we consider the many-good nonlinear tax model. We focus on different versions of such a model depending on assumptions on individuals' decisions on savings and labour supply, i.e. individuals' time-consistency over their lifetime.

In the Mirrlees (1971) model the rich are different from the poor in only one way; they are endowed with the ability to command a higher market wage rate, which is assumed to reflect the higher productivity of their labour effort. In fact, there is a variety of other reasons to why some people end up affluent and others do not. One might well argue that both high-income and low-income people do not owe their (un)success entirely to ability, but that some of the income differentials are due to luck, different time preferences and inherited resources.[3] Hence, unlike in the original Mirrlees model, we assume that individuals do not differ only in ability, but also in time preferences. It is quite plausible to assume that in reality, for all sorts of reasons, both ability and time preferences are not observable. Multidimensional heterogeneity in agents' characteristics is a realistic assumption but it complicates the analysis notably, as discussed e.g. in Boadway et al. (2002). There are some analytical studies in a discrete case with two-dimensional heterogeneity,[4] but they are simplified further to a three-type case

commitment device. Roberts (1984) and Gaube (2007) do not, however, consider taxes on savings as possible tax instruments. See also Berliant and Ledyard (2005) for a recent contribution.

[3] There are also considerable variations in the rate of return people receive on their capital. It is quite possible that some of variations in the return on capital are the result of the application of skill and effort; but most is clearly the result of risky outcomes (luck).

[4] An exception is Tuomala and Tarkiainen (1999; 2007), who present numerical calculations in the two-dimensional optimal tax problem for a continuum of agents.

(Cuff, 2000; Blomquist and Christiansen, 2004). Cremer, Lozachmeur and Pestieau (2004) also consider social security and retirement decision in a static model with agents differing in two dimensions, productivity and health status. Moreover, due to the nature of the heterogeneities they use, the pattern of the constraints for optimal tax system is simpler than in the general case.

Economists such as Wicksell, Pigou, Ramsey and Allais[5] have argued for the case that the social planner should be more patient than individuals. For example, Ramsey (1928) claimed that it was "ethically indefensible" not to take into account discounting of the future utilities. Any such argument can be called paternalistic. The notion that individuals may not make the best choices for themselves raises difficult issues. Individuals may be fully rational and they just happen to have a high preference for the present, which causes them to save little, because too little weight is given to future contingencies. Should the government be welfarist and maximise individual welfare, as the individual sees it? Or, as suggested in recent behavioural public economics literature, should it be paternalistic, or non-welfarist, and discount the future at a different rate than individuals? In fact, individuals may want the government to intervene, to induce behaviour that is closer to what individuals wish they were doing. For example, in models with quasi-hyperbolic preferences it is typically desirable to impose a particular savings plan on individuals.

[5] "In actual societies it seems to be common that social choices deviate from consumer preferences in the assessment of the relative importance of future needs with respect to present needs…Public saving and legal arrangements such as compulsory pension schemes allow this objective to be realised. …It was in order to generalise optimum theory to such a collective attitude that M. Allais put forward the concept of ´rendement social généralisé´. His idea is to define and investigate a notion of optimum in which individual preferences are retained for the choice between consumption relating to the same date, but not necessarily between those relating to different dates." (Malinvaud, 1972, p. 244)

The object of this paper is to characterise the optimal lifetime redistribution policy within a cohort. We consider the optimal tax treatment of savings when individuals differ both in abilities and time preferences both in welfarist and paternalistic cases.[6] The case with two-dimensional heterogeneity is considered both in analytical form, and, to gain a better understanding, also in the light of numerical examples. A numerical solution allows us also to consider the dependence of the replacement rate on earnings and to compare consumption dispersion in both periods. We also contemplate how an extreme form of time consistency, myopia, affects the results.

The structure of the paper is as follows. Section 2 considers a two-period and two-type version of the Mirrlees model with a positive correlation between skill and discount factor with both paternalistic and welfarist governments. In Section 3 the assumption of perfect correlation is given up and the model is generalised into a three-type case. The optimal lifetime redistribution is considered both with analytical and numerical tools. A case with myopic individuals is also examined. In Section 4 we deliberate a general four-type model with help of a numerical approach. Section 5 concludes.

2. A benchmark model: two types with a positive correlation between skill and discount factor

Unlike the original Mirrlees model we assume that individuals do not differ only in productivity, but also in discount rate or time preference[7]. As a benchmark we

[6] In a longer working paper version Tenhunen and Tuomala (2007) a case of Rawlsian government is also considered.

[7] Sandmo (1993) considers a case where people differ in preferences, but are endowed with the same resources. Tarkiainen and Tuomala (1999, 2007) also consider a continuum of taxpayers simultaneously distributed by skill and preferences for leisure and income.

use a simple two-type model, similar to the much used two-type model first introduced by Stern (1982) and Stiglitz (1982). Each individual has a skill level reflecting his wage rate, denoted by *n*, and a discount factor, denoted by δ. We denote low-skilled types by the superscript L and high-skilled types by the superscript H. The assumption of positive correlation implies that $\delta^L \leq \delta^H$. The proportion of individuals of type i in the population is $N^i > 0$, with $\sum N^i = 1$.

As is well known, due to Atkinson and Stiglitz (1976), under a mild separability assumption, income taxation does not need be supplemented by other taxes. Saez (2002) argues that the Atkinson-Stiglitz result of commodity taxes holds when each individual has identical discount rates. He also argues that individuals with higher earnings save relatively more, which suggests that high-skilled individuals are more likely to have higher discount factors. In this case, discount factor is positively correlated with productivity level. For this reason we take as a starting point separable utility with different time preference.[8] The life-time utility of an individual of type i is additive in the following way:

$$U^i = u(c^i) + \delta^i v(x^i) + \psi(1 - y^i), \qquad (1)$$

where *c* and *x* denote respectively consumption when young and when retired, and *y* is labour supply when young. It is increasing in each argument and $u', v', \psi' > 0$ and $u'', v'', \psi'' < 0$ and strictly concave. We also assume that all goods are normal.

To introduce return to capital and the possible taxation thereof, it is necessary to consider a two-period model. Individuals are free to divide their first period income between consumption, *c* and savings, *s*. Each unit of savings yields a

[8] Alternatively the same outcome could be reached by assuming homothetic preferences and linear Engel curves.

consumer an additional 1+θ units of consumption in the second period after tax income, x. As a further simplification we assume that the return to savings, θ, is fixed, which may be justified by assuming that we consider a small open economy facing a world capital market. Consumption in each period is given by $c^i = n^i y^i - T(n^i y^i) - s^i$ and $x^i = (1+\theta)s^i$, i = L,H.

The government wishes to design a lifetime tax system that may redistribute income between individuals in the same cohort. There is asymmetric information in the sense that the tax authority is informed neither about individual skill levels, labour supply nor discount rates. It can only observe before–tax income, ny. In this setting, where taxes both on earnings and savings income are available we examine whether or not savings ought to be taxed. The separability assumption makes it possible to isolate the significance of variations in time preferences.

2.1. The welfarist government

First we analyse a model in which the government respects the individual sovereignty principle and evaluates individuals' well-being using their own discount rates. Assume that the government controls c^i, x^i and y^i directly. Alternatively, if we assume that there are no private savings we have a model of labour income taxation in the first period and public provision of pension in the second period.

In the welfarist case government's problem is to maximise the following social welfare function

$$\sum N^i (u(c^i) + \delta^i v(x^i) + \psi(1 - y^i)) \tag{2}$$

subject to the revenue constraint

$$\sum N^i\left(n^i y^i - c^i - rx^i\right) = R \tag{3}$$

where $r = \dfrac{1}{1+\theta}$, and the self-selection constraint[9]

$$u\left(c^H\right) + \delta^H v\left(x^H\right) + \psi\left(1 - y^H\right) \geq \hat{u}\left(c^L\right) + \delta^H \hat{v}\left(x^L\right) + \hat{\psi}\left(1 - \frac{n^L}{n^H} y^L\right) \tag{4}$$

where the terms with a 'hat' refer to mimicking behaviour. Multipliers λ and μ are attached respectively to the budget constraint and the self-selection constraint. The Lagrange function of the optimization problem is

$$\begin{aligned} L = \sum_i N^i\left[u\left(c^i\right) + \delta^i v\left(x^i\right) + \psi\left(1 - y^i\right)\right] + \lambda\left[\sum N^i\left(n^i y^i - c^i - rx^i\right) - R\right] \\ + \mu\left[u\left(c^H\right) + \delta^H v\left(x^H\right) + \psi\left(1 - y^H\right) - \hat{u}\left(c^L\right) - \delta^H \hat{v}\left(x^L\right) - \hat{\psi}\left(1 - \frac{n^L y^L}{n^H}\right)\right] \end{aligned} \tag{5}$$

As has become conventional in the literature we may interpret the marginal rate of substitution between gross and net income as one minus the marginal income tax, $\dfrac{\psi'(ny/n)}{nu_c} = 1 - T(ny)$, which would be equivalent to the characterisation of the labour supply of an agent facing an income tax function $T(ny)$. The marginal labour income tax rates satisfy the usual properties; $T'(n^L y^L) > 0$ and $T'(n^H y^H) = 0$.

[9] The direction of the binding self-selection constraint is assumed to be, following the tradition in the one-dimensional two-type model, from high-skilled individual towards low-skilled individual. This pattern is also confirmed by the numerical simulations.

Our main interest is in the marginal taxation of savings. For this purpose the first order conditions are written in the form $\left(\frac{u_c}{v_x}\right)^i = \frac{\delta^i}{r}\left[1 - d^i\right]$, where the left hand side is individual i's marginal rate of substitution between consumption in the first and in the second period and d^i is the distortion. A positive (negative) d^i implies that type i should have an implicit tax (subsidy) on savings. It is useful to define a relative difference in discount factors as $\Delta^{ij} \equiv \frac{\delta^i - \delta^j}{\delta^j}$, for any pair of discount factors. The first order conditions (Appendix A) imply that

$$\begin{aligned} d^L &= \left(\varphi^1 - 1\right)\Delta^{HL} \\ d^H &= 0 \end{aligned} \tag{6}$$

where $\varphi^1 = \frac{N^L}{N^L - \mu}$. The returns to savings of type i should not be taxed when d^i is zero. With d^H=0, the optimal implicit marginal tax rate for the high-skill type is zero. When we assume, empirically plausibly, $\delta^H > \delta^L$, we have $d^L>0$ implying implicit taxation of savings for the low-skill type. This is the same result as in Diamond (2003).

As a result of the two-dimensional heterogeneity, a tax on capital income is an effective way to relax an otherwise binding self-selection constraint. This is because even under separability the mimicker and the individual mimicked do not save the same amount. A high-skilled individual choosing to mimic someone with less skill values savings more than a low-skilled individual, since discount of the future is less for the potential mimicker. Thus, taxing savings relaxes the self-selection constraint. Or put in another way: distortions generate second-order efficiency costs but first-order redistributional benefits.

2.2. Differences in government's and individuals' discount rates

Using a sum of utilities as a social objective function may entail some ethical objections. Namely, if individuals are identical in preferences, equal marginal utilities of all coincide with equal total utilities. Sen (1973) pointed out that with the diversity of human beings (different time preferences in our case) the two may pull in the opposite directions. For those who prefer to think of the justification for redistribution as being based on the inequality of opportunity, differences in preferences may provide a suitable basis for distinguishing economic rewards but differences in skills in turn do not. This point of view raises questions on the nature of the parameter δ. It may be argued that both skills and time preferences are "circumstances of birth". Therefore, it is far from clear how suitable the utilitarian criterion is in this case. These considerations may be one reason to also consider paternalistic social objectives.

Assume now that the government does not respect the individual sovereignty principle and evaluates individuals' well-being using a discount rate different from those of individuals. Now the paternalistic government with a discount factor δ^g maximizes social objective given by

$$\sum N^i (u(c^i) + \delta^g v(x^i) + \psi(1 - y^i)) \tag{7}$$

In the first-best situation there would be a subsidy on savings, equal to Δ^{gi}, that exactly corrects the difference in discount factors. In the second-best case with non-linear taxation the outcomes generated by individual preferences enter government's budget constraint and thus they also have to be taken into account.

The form giving the implicit marginal taxes for savings is now $\left(\frac{u_c}{v_x}\right)^i = \frac{\delta^g}{r}\left[1 - \alpha^i\right]$, where α^i gives the distortion. However, from the perspective of the individual the distortion is still given by d^i in $\left(\frac{u_c}{v_x}\right)^i = \frac{\delta^i}{r}\left[1-d^i\right]$. There is a correspondence between these two distortions, given by $\alpha^i = d^i + \Delta^{ig}\left(d^i - 1\right)$, for all types *i*. It can be observed, that as long as $d^i < 1$ (i.e. the implicit distortion for savings from individuals' point of view is less than 100 per cent), $\alpha^i > d^i$. In other words, a positive (negative) α^i implies that the savings decision of type *i* is under-subsidized (over-subsidized) relative to the first-best case from individuals' point of view.[10]

From the first-order conditions of government's problem (Appendix A) we get the following expressions:

$$\begin{aligned} \alpha^L &= \left(\varphi^1 - 1\right)\Delta^{Hg} \\ \alpha^H &= \left(\varphi^2 - 1\right)\Delta^{Hg} \end{aligned} \tag{8}$$

where $\varphi^1 = \frac{N^L}{N^L - \mu}$ and $\varphi^2 = \frac{N^H}{N^H + \mu}$.

From Eq. (8) we see that the returns to savings should be distorted: α^L is negative and α^H is positive, when type H individual's discount factor is smaller than that of the social planner. In a special case, where $\delta^H = \delta^g$, α^H reduces to zero. This implies the following proposition[11]

Proposition 1: *As long as $\delta^H < \delta^g$, the savings of the high-skill type should be implicitly taxed at the margin (under-subsidized relative to the first-best situation) while the savings of the low-skill type should be implicitly subsidized (over-subsidized relative to the first-best situation).*

---

[10] In the first best situation the distortion is $-\Delta^{gi}$, which corrects the effect of the difference in the discount rates.

[11] We are grateful to Thomas Gaube for suggesting this interpretation.

Relaxing the self-selection constraint in this paternalistic case means an upward distortion for the savings of type L and a downward distortion for the savings of type H. Contrary to the welfarist case, the savings of the low-skill type are now subsidized. In other words, compared to the savings required to the first best optimum, it is socially beneficial to have the high-skill type save less and the low-skill type save more. Proposition 1 implies that non-linear taxes and subsidies of savings are a useful complement to non-linear taxation of labour income for a government interested in redistribution.

2.3 Numerical simulations

To investigate further the properties of the optimal redistribution we now turn to numerical examples. We assume the following separable forms of utility functions; $U^i = \log c^i + \delta^i \log x^i + \log(1 - y^i)$ (CD) and $U^i = -\frac{1}{c^i} - \delta^i \frac{1}{x^i} - \frac{1}{1 - y^i}$ (CES). We choose the following parameterization:

Fraction of individuals in each group	$N^i = 0.5$ for i = L, H
Discount factors	$\delta^L = 0.6$, $\delta^H = 0.8$, $\delta^g = 1$, r = 0.95
Productivities (wages)	$n^L = 2$, $n^H = 3$

TABLE 1: Parameterisation in the two-type economy

No a priori assumptions of the binding self-selection constraints are made in numerical simulations. The results verify the assumption that the only binding self-selection constraint is type H considering mimicking type L[12].

Numerical results allow us to consider the overall dispersion of consumption. One way is to consider standard Gini-coefficients in both periods respectively. It turns out that we can also make welfare rankings between the distributions of the first and the second period consumption based on the results by Atkinson (1970) and Shorrocks (1980). They determine welfare ranking with the help of Lorenz dominance which requires less normative assumptions on the social welfare function. As a corollary of Atkinson's theorem we know that with different mean incomes, unambiguous welfare ranking survives only when it is the Lorenz-dominant distribution that has the higher means, i.e. welfare is greater at the period j when mean consumption is greater and the Lorenz curve L_j is inside the Lorenz curve of the other period. Shorrocks' result can be applied when Lorenz curves cross or when a Lorenz-dominant distribution has a lower mean. The result uses the generalised Lorenz curve of period j, GL_j, for comparison; higher GL in all points implies higher welfare.

Tables 2a and 2b present the numerical results for the welfarist and paternalist cases: private utilities in the optimum, the values of the marginal rates of labour income T', the marginal tax rates of savings, given by distortion d^i in the welfarist case and α^i in the paternalistic case, and the replacement rates in terms of second period consumption relative to the first period gross income (x/ny)[13].

[12] The slackness of the other self-selection constraints is also checked by calculating the difference in utilities when mimicking and when not.

[13] Further details on the numerical simulations are provided in Appendix B, Tables B1 for welfarist case and Tables B2 for paternalistic case.

		U	T'	d	x/ny	Consumption dispersion
CD	type L	-1.43	3.56	3.69	44.61	means: $\bar{c}$ = 0.92 $\bar{x}$ = 0.68 Lorenz dominance: $L_c > L_x$ Gini coefficients: G_c= 0.050 G_x= 0.128
	type H	-1.21	0	0	42.62	
CES	type L	-4.67	6.30	6.72	51.20	means: $\bar{c}$ = 0.73 $\bar{x}$ = 0.62 Lorenz dominance: $L_c > L_x$ Gini coefficients: G_c= 0.042 G_x= 0.086
	type H	-4.56	0	0	44.42	

TABLE 2a. The welfarist case: utility levels, marginal tax rates on labour income and saving (per cent), replacement rates and consumption dispersion

		U	T'	α	x/ny	Consumption dispersion
CD	type L	-1.55	6.83	-4.40	59.33	means: $\bar{c}$ = 0.82 $\bar{x}$ = 0.86 Lorenz dominance: $L_c < L_x$ Gini coefficients: G_c= 0.090 G_x= 0.072
	type H	-1.18	0	3.05	48.54	
CES	type L	-4.79	11.07	-7.47	58.87	means: $\bar{c}$ = 0.69 $\bar{x}$ = 0.71 Lorenz dominance: $L_c < L_x$ Gini coefficients: G_c= 0.069 G_x= 0.055
	type H	-4.52	0	4.28	47.62	

TABLE 2b. The paternalistic case: utility levels, marginal tax rates on labour income and saving, replacement rates and consumption dispersion

The analytical results on the marginal tax rates are confirmed by the numerical simulations. The results also suggest that the retirement consumption, measured in terms of Gini coefficient, is less dispersed than the first period consumption in the paternalistic case, whereas in the welfarist case the ordering is reversed. It turns out that in the welfarist case the distribution of the first period consumption is the Lorenz-dominant distribution with a higher mean. In the paternalist case the dominance is the other way round. According to the numerical results the replacement rate (x/ny) decreases in earnings. Moreover, our numerical results suggest that the paternalistic government policy increases second period consumption compared to the welfarist government policy (see Appendix B, Tables B1-B2).

3. A three-type case

In this section we generalise the previous model by abandoning the assumption that productivity and time preference are perfectly correlated. In general, there are now four types of individuals who differ both in productivity and time preference numbered as follows (Table 3):[14]

	low-skilled	**high-skilled**
low delta	type 1	type 3
high delta	type 2	type 4

TABLE 3: Types of individuals

To maintain the tractability, we simplify the model further by assuming that there are actually only three types: low-skilled types with a low discount factor δ^L, indexed as type 1, high-skilled types with a low discount factor, type 3, and high-skilled types with a high discount factor, type 4. [15] Assume also that utility is given by Eq. (1), the same additively separable form as in the two-type case.

In the three-type case there are several possibilities for mimicking. When the self-selection constraints bind depends now on how the labour supply and the

[14] Despite the numbering it is here chosen to follow the case with one-dimensional heterogeneity, where types are ordered usually with respect to their income, consumption or utilities, in a two-dimensional world the ordering is not self-evidently clear.

[15] There are also other possible structures for a three-type economy. A discussion on the issue is presented in Appendix C.

distributional preferences of the government hinge on the time preferences and skill level. Without any assumptions of the mimicking behaviour there are six possible self-selection constraints given by

$$u\left(c^i\right)+\delta^i v\left(x^i\right)+\psi\left(1-y^i\right)\geq \hat{u}^{ij}\left(c^j\right)+\delta^i\hat{v}^{ij}\left(x^j\right)+\hat{\psi}^{ij}\left(1-\frac{n^j y^j}{n^i}\right) \tag{9}$$

for i,j =1,3,4 and $i \neq j$

It is typical in the multidimensional screening problems to focus on the so-called relaxed problem. In this case, the relaxed problem is defined as the problem where only the three downward incentive constraints are considered. Of course, the solution of the relaxed problem is of no intrinsic interest except insofar as it yields the solution to the fully constrained problem. In general, other self-selection constraints beside downward constraints may also be binding at the optimum. In those cases the relaxed problem is no longer relevant. Moreover, it is not easy to ascertain the direction of redistribution. In cases where the relaxed problem is not the solution of the full problem, the 'no distortion at the top' result no longer holds.

Analytically it is not possible to determine which of the self-selection constraints are binding. We solve the problem first numerically to determine the binding self-selection constraints. The utility functions are the same as in the two-type case; Cobb-Douglas and CES forms. The parameters used are otherwise the same as before, except for the distribution of types, which is now given by N^1 = 0.5, N^3 = 0.25, N^4 = 0.25. On the basis of numerical simulations we choose the following pattern of binding self-selection constraints ((4,3) (3,1) and (4,1))[16] given by:

[16] Interestingly, not only in simulations displayed in Tables 4 and 5 but also in many other simulations, this pattern of selection constraints appears systematically.

$$u\left(c^4\right)+\delta^H v\left(x^4\right)+\psi\left(1-y^4\right)\geq \hat{u}^{43}\left(c^3\right)+\delta^H \hat{v}^{43}\left(x^3\right)+\hat{\psi}^{43}\left(1-y^3\right) \tag{10a}$$

$$u\left(c^3\right)+\delta^L v\left(x^3\right)+\psi\left(1-y^3\right)\geq \hat{u}^{31}\left(c^1\right)+\delta^L \hat{v}^{31}\left(x^1\right)+\hat{\psi}^{31}\left(1-\frac{n^1 y^1}{n^3}\right) \tag{10b}$$

$$u\left(c^4\right)+\delta^H v\left(x^4\right)+\psi\left(1-y^4\right)\geq \hat{u}^{41}\left(c^1\right)+\delta^H \hat{v}^{41}\left(x^1\right)+\hat{\psi}^{41}\left(1-\frac{n^1 y^1}{n^4}\right) \tag{10c}$$

When it turns out that only the constraints (10a) and (10b) are binding in the optimum the economy ends up in a separating equilibrium. The other possibility is a bunching optimum, where at least two of the types are indistinguishable. Bunching at the bottom is not an interesting case here as it leads back to the traditional two-type case analysed thoroughly in the earlier literature. Bunching at the top, where the high-skilled types 3 and 4 are indistinguishable is also a possible equilibrium that corresponds to the two-type case with perfect correlation. Thus, we concentrate here only on the separating equilibrium.

3.1 The welfarist social planner

The welfarist government maximises a sum of utilities (2) subject to the revenue constraint (3) and the self-selection constraints (9). According to the numerical simulations, the binding self-selection constraints turns out to be (10a) and (10b). Using this information, the Lagrange function of the optimisation problem is

$$\begin{aligned} L = \sum N^i \left[u\left(c^i\right) + \delta^i v\left(x^i\right) + \psi\left(1 - y^i\right)\right] \ + \lambda\left[\sum N^i\left(n^i y^i - c^i - r x^i\right) - R\right] \\ + \mu^{43}\left[u\left(c^4\right) + \delta^H v\left(x^4\right) + \psi\left(1 - y^4\right) - \hat{u}^{43}\left(c^3\right) - \delta^H \hat{v}^{43}\left(x^3\right) - \hat{\psi}^{43}\left(1 - y^3\right)\right] \\ + \mu^{31}\left[u\left(c^3\right) + \delta^L v\left(x^3\right) + \psi\left(1 - y^3\right) - \hat{u}^{31}\left(c^1\right) - \delta^L \hat{v}^{31}\left(x^1\right) - \hat{\psi}^{31}\left(1 - \frac{n^1 y^1}{n^3}\right)\right] \end{aligned} \qquad (11)$$

where multipliers λ and μ are attached to the budget constraint and the binding self-selection constraints respectively.

Following the analysis in the two-type case, we can rewrite the first order conditions (presented in Appendix A) as $\left(\frac{u_c}{v_x}\right)^i = \frac{\delta^i}{r}\left[1 - d^i\right], i = 1,3,4$, where the discount factors are $\delta^1 = \delta^3 = \delta^L$ and $\delta^4 = \delta^H$, and the distortions are given by d^i:

$$\begin{aligned} d^1 &= 0 \\ d^3 &= \frac{\mu^{43}}{N^3 - \mu^{43} + \mu^{31}} \Delta^{HL} \\ d^4 &= 0 \end{aligned} \qquad (12)$$

where $\Delta^{HL} = \frac{\delta^H - \delta^L}{\delta^L} > 0$. Equation (12) implies the following proposition

Proposition 2. *The savings decisions of type 1 and type 4 are not distorted, and hence not taxed at the margin. Type 3 faces a positive marginal tax rate on savings (under-subsidized relative to the first-best).*

Proposition 2 implies that even when the government respects consumers' time preferences, there is a distortion for type 3. This results from the fact that in a model with the given binding self-selection constraints (3,1) and (4,3) a tax on capital income can be used to mitigate otherwise binding self-selection constraints.

The numerical solution in Table 4, giving the utility levels, the marginal tax rates on income and savings, the replacement rates and information on consumption dispersion, present a number of interesting features. For example, the constraint (4,3) binds even though type 3 has a higher utility than type 4.[17] Hence, we cannot extend the intuition based on the one-dimensional model to the two-dimensional case. We shall return this question in the context of a four-type model in Section 4. The numerical results confirm that savings of type 3 are taxed at the margin. The replacement rates are non-monotonic; the replacement rate for type 3 is lower than those of types 1 and 4. The result holds for both types of utility functions.

The Gini coefficients for inequality in first period consumption and retirement consumption in Table 4 show that consumption is less dispersed in the first period than in the second period. This seems plausible, as the welfarist government does not try to correct the time preference for the second period consumption. Allowing greater inequality in consumption levels is a result of respecting the sovereignty of consumers. In this case the distribution of the first period consumption is the Lorenz-dominant distribution that has a higher mean.

		U	T'	d	x/ny	Consumption dispersion
CD	type 1	-1.48	5.32	0	44.39	means: $\bar{c}$ = 0.97 $\bar{x}$ = 0.68 Lorenz dominance: $L_c > L_x$ Gini coefficients: G_c= 0.081 G_x= 0.114
	type 3	-1.12	0	2.59	36.21	
	type 4	-1.19	0	0	43.14	
CES	type 1	-4.73	8.87	0	51.00	means: $\bar{c}$ = 0.76 $\bar{x}$ = 0.63 Lorenz dominance: $L_c > L_x$ Gini coefficients: G_c= 0.061 G_x= 0.077
	type 3	-4.22	0	4.52	41.26	
	type 4	-4.53	0	0	44.97	

TABLE 4: The numerical solution in the welfarist case. Binding self-selection constraints are **(3,1)** and **(4,3)**.

[17] This is not necessarily surprising in multi-dimensional problems; see for example Judd and Su (2006) and Cremer et al. (2001).

3.2 Differences in government's and individuals' discount rates

Next we consider a case in which the government has a discount factor δ^g, i.e. it maximises (7) subject to the revenue constraint (3) and the self-selection constraints (9). With the same parameterisation as before the numerical simulation shows that in the optimum only constraints (10a) and (10b) are binding and the economy ends up in a separating equilibrium. The Lagrange function is now given by

$$\begin{aligned} L = & \sum N^i \left[u\left(c^i\right) + \delta^g v\left(x^i\right) + \psi\left(1 - y^i\right)\right] + \lambda\left[\sum N^i \left(n^i y^i - c^i - r x^i\right) - R\right] \\ & + \mu^{43}\left[u\left(c^4\right) + \delta^H v\left(x^4\right) + \psi\left(1 - y^4\right) - \hat{u}^{43}\left(c^3\right) - \delta^H \hat{v}^{43}\left(x^3\right) - \hat{\psi}^{43}\left(1 - y^3\right)\right] \\ & + \mu^{31}\left[u\left(c^3\right) + \delta^L v\left(x^3\right) + \psi\left(1 - y^3\right) - \hat{u}^{31}\left(c^1\right) - \delta^L \hat{v}^{31}\left(x^1\right) - \hat{\psi}^{31}\left(1 - \frac{n^1 y^1}{n^3}\right)\right] \end{aligned} \tag{13}$$

We can express the first order conditions (given in Appendix A) from government's perspective as $\left(\frac{u_c}{v_x}\right)^i = \frac{\delta^g}{r}\left[1 - \alpha^i\right], i = 1,3,4$. The distortions government observes are given by

$$\begin{aligned} \alpha^1 &= \left(\varphi^1 - 1\right)\Delta^{Lg} \\ \alpha^3 &= \frac{\mu^{43}}{N^3 - \mu^{43} + \mu^{31}}\Delta^{Hg} - \frac{\mu^{31}}{N^3 - \mu^{43} + \mu^{31}}\Delta^{Lg} \\ \alpha^4 &= \left(\varphi^4 - 1\right)\Delta^{Hg} \end{aligned} \tag{14}$$

where $\varphi^1 = \frac{N^1}{N^1 - \mu^{31}}$ and $\varphi^4 = \frac{N^4}{N^4 + \mu^{43}}$.

Now there are terms resulting from both the paternalist objectives (Δ^{Lg}, Δ^{Hg}) and from the distributional considerations (terms including μ^{31}, μ^{43}). They cannot, however, be separated to isolate the effects of these two parts. Even when without paternalistic objectives types 1 and 3 were undistorted, in the case with paternalistic government the optimal distortions for these types depend on both effects. In (14) terms with Δ^{Lg} and Δ^{Hg} are negative as long as the social planner has a higher discount factor than types 3 and 4 with δ^{H}. These results give rise to the following proposition.

Proposition 3. *As long as $\delta^{g} > \delta^{H} \left(> \delta^{L}\right)$, for type 1 the marginal taxation of saving is negative (over-subsidized relative to the first best) and the marginal tax on savings for type 4 is positive (under-subsidized relative to the first best). For type 3 the sign of the marginal rate is indeterminate.*

There are two distortions with opposite signs for type 3, so the overall effect on the tax on savings is ambiguous. Our numerical solution implies that the optimal savings tax rate for type 3 is positive, i.e. there is an implicit tax on savings (Table 5). For type 1 there is an implicit subsidy and for type 4 a tax, as also suggested by the analytical results. The tax for type 3 seems to be systematically larger than that for type 4. Note that the marginal subsidy for type 1 and the marginal tax for type 4 also contribute to the objective of the government (paternalism).

The replacement rates show a similar non-monotonic pattern as in the welfarist case: type 3 has the lowest replacement rates. The dispersion in consumption in both periods is now reversed compared to the welfarist case: second period consumption is now more equally distributed than consumption in the first period. The Lorenz dominance also supports the view as the second period is Lorenz-dominant to the first period.

		U	T’	α	x/ny	Consumption dispersion
CD	type 1	-1.59	7.95	-10.37	60.11	means: $\bar{c}$ = 0.87 $\bar{x}$ = 0.90 Lorenz dominance: $L_c < L_x$ Gini coefficients: G_c=0.105 G_x=0.064
	type 3	-1.14	0	10.35	46.92	
	type 4	-1.15	0	3.23	49.02	
CES	type 1	-4.84	12.81	-17.63	59.58	means: $\bar{c}$ = 0.72 $\bar{x}$ = 0.73 Lorenz dominance: $L_c < L_x$ Gini coefficients: G_c= 0.080 G_x= 0.048
	type 3	-4.23	0	14.13	46.54	
	type 4	-4.50	0	4.49	47.99	

TABLE 5: The numerical solution in the paternalistic case, binding self-selection constraints (3,1) and (4,3)

It is clear that as long as $\delta^g > \delta^H \left(> \delta^L\right)$ and the paternalistic government values the second period consumption of each type with a common discount factor, the tax system is designed so that the resulting consumption dispersion in the second period is more equal than with a government respecting consumers’ own time preferences. This is in accordance with Diamond (2003), who analyses the dependence between replacement rates (second period consumption relative to first period consumption) and risk aversion. He finds that when the elderly are more risk averse than younger people the optimal lifetime redistribution tends to imply that retirement consumption should be less dispersed than first period consumption. In our case the idea of the risk of having low consumption in the retirement period is internalised by the paternalist government.

3.3 Myopic behaviour

Next we analyse the model in which not all individuals save voluntarily and in their labour supply decisions they ignore the implications of their earnings when young for the retirement income, i.e. they are myopic. Myopic behaviour may be quite common for a substantial proportion of individuals who hardly save at all and rely almost entirely on public pension benefits. Does this model generate less dispersed retirement consumption than the models with rational behaviour? Diamond (2003, Chapter 4) and Diamond and Mirrlees (2000) consider a benchmark situation where individuals do not save at all. In their model workers are otherwise identical, but their skills differ, and the government's objective is to design an optimal redistributive policy for those of working age and for the retired. If the social welfare function exhibits inequality aversion, the optimal retirement consumption is shown to be higher for those whose lifetime income has been smaller.

It is worth stressing that myopic behaviour is distinct from the behaviour associated with heavy discounting of the future. If individuals have low discount factors, they will save little for their retirement consumption, but this reflects optimising behaviour. By contrast, if individuals are myopic, and are subjected to compulsory saving, their welfare will increase. Although the behavioural foundations of myopia differ essentially from those of time consistent utility maximisation, the analysis developed above can be used with minor modifications. We simply interpret that the discount factor δ being either 0 for perfectly myopic types and 1 for completely rational individuals.

Myopic labour supply implies that the retirement consumption does not enter the incentive compatibility constraint of a myopic mimicker. The social welfare function depends on *ex post* utilities given by $u(c^i)+\delta^g v(x^i)+\psi(1-y^i)$.

By assuming that all types are myopic, we get a two-type version of the continuum model analysed in Diamond (2003). In this case the basic justification for a pension system is to guarantee some level of resources in the retirement period. A more interesting case is to extend Diamonds' analysis to a more realistic case in which some people save and some do not, i.e. assume that myopia and ability are imperfectly correlated. We assume that $\delta^L = 0$, but contrast to Diamond (2003), we allow social planner and type 4 individuals to have different discount factors.

Myopic types perceive only the apparent utility $u(c^i)+\psi(1-y^i)$. Thus there is now no first-best case, where the government could by a subsidy on savings induce myopic types to save voluntarily.[18] As a result, the interpretation based on private distortion *d* cannot be used here. However, in the optimum, from government's point of view, there is also a distortion α for myopic types.

Suppose now that myopic behaviour is illustrated by δ^L being zero, i.e. types 1 and 3 are myopic whereas type 4 is not. Myopic types take only the apparent utility into account, also when mimicking. But rational type 4 perceives the changes in second period consumption that would occur if he mimicked myopic types, so the part reflecting the utility from the second period consumption remains in the self-selection constraints binding type 4 but vanishes from those binding types 1 and 3.

The social planner maximises (7) subject to the revenue constraint (3) and the self-selection constraints (9). Numerical simulation suggests that all he self-selection constraints (10a), (10b) and (10c) are binding.

The Lagrange function for the problem can be written as

$$\begin{aligned} L = &\sum N^i\left[u(c^i)+\delta^g v(x^i)+\psi(1-y^i)\right] + \lambda\left[\sum N^i\left(n^i y^i - c^i - rx^i\right)-R\right] \\ &+\mu^{43}\left[u(c^4)+\delta^H v(x^4)+\psi(1-y^4)-\hat{u}^{43}(c^3)-\delta^H\hat{v}^{43}(x^3)-\hat{\psi}^{43}(1-y^3)\right] \\ &+\mu^{41}\left[u(c^4)+\delta^H v(x^4)+\psi(1-y^4)-\hat{u}^{41}(c^1)-\delta^H\hat{v}^{41}(x^1)-\hat{\psi}^{41}\left(1-\frac{n^L y^1}{n^H}\right)\right] \quad (15) \\ &+\mu^{31}\left[u(c^3)+\psi(1-y^3)-\hat{u}^{32}(c^1)-\hat{\psi}^{31}\left(1-\frac{n^L y^1}{n^H}\right)\right] \end{aligned}$$

---

[18] The first-best distortion $-\frac{\delta^g-\delta^i}{\delta^i}$ approaches $-\infty$ when δ^i goes to zero.

Writing the first order conditions in a short form as $\left(\frac{u_c}{v_x}\right)^i = \frac{\delta^g}{r}\left[1-\alpha^i\right]$, $i = 1,3,4$ gives us the distortions

$$
\begin{aligned}
\alpha^1 &= \frac{\mu^{41}}{N^1 - \mu^{41} - \mu^{31}}\Delta^{Hg} - \frac{\mu^{31}}{N^1 - \mu^{41} - \mu^{31}} \\
\alpha^3 &= \frac{\mu^{43}}{N^3 - \mu^{43} + \mu^{31}}\Delta^{Hg} + \frac{\mu^{31}}{N^3 - \mu^{43} + \mu^{31}} \\
\alpha^4 &= -\frac{\mu^{43} + \mu^{41}}{N^4 + \mu^{43} + \mu^{41}}\Delta^{Hg}
\end{aligned}
\tag{16}
$$

The analytical results follow the earlier case. For type 1 the marginal taxation of saving is negative, for type 4 positive, and the sign of the distortion is indeterminate for type 3.

The numerical solution (Table 6) with otherwise same parameters as in the earlier case, except for δ^L=0, also follows the same lines as the paternalistic case presented earlier. The optimal savings tax rate for type 1 is negative, i.e. there is an implicit subsidy for savings as suggested by the analytical part. For types 3 and 4 we have a tax.

The non-monotonicity of the replacement rates is reinforced in the case with myopia. Following the paternalistic case without myopia, the dispersion in the first period consumption is also greater than in the second period. In point of fact, the decrease in the Gini coefficient from the first to the second period is steeper than in a case without myopia. In this case both Lorenz curves (L) and generalised Lorenz curves (GL) cross. The mean consumption in both periods is also practically the same. Hence we cannot make unambiguous welfare rankings between distributions.

When we interpret the model as a public pension system with no private savings, the previous result indicates that the pension system is more generous to low-skilled types and less generous to high-skilled types in the economy with myopia. Accordingly, comparing Gini-coefficients in paternalistic cases with and without myopia suggests that the pension system with myopia is more redistributive than in the case without myopia. Finally, as in a two-type model the paternalistic government policy increases the second period consumption compared to the welfarist government policy (see Appendix B, Tables B4-B5 for further numerical results).

		U	T'	α	x/ny	Consumption dispersion
CD	type 1	-1.47	9.15	-23.77	64.53	means: $\bar{c}$ = 0.88 $\bar{x}$ = 0.87 Lorenz dominance: L & GL curves cross Gini coefficients: G_c=0.124 G_x=0.031
	type 3	-1.02	0	25.65	41.71	
	type 4	-1.18	0	3.13	48.35	
CES	type 1	-3.96	14.61	-40.07	62.86	means: $\bar{c}$ = 0.72 $\bar{x}$ = 0.71 Lorenz dominance: L & GL curves cross Gini coefficients: G_c=0.092 G_x=0.023
	type 3	-3.35	0	33.85	42.90	
	type 4	-4.52	0	4.37	47.48	

TABLE 6: The numerical solution in the paternalistic case with myopia, binding self-selection constraints (3,1), (4,1) and (4,3)

In sum, Tables 4 - 6 show that the replacement rates for type 3 are lower than those of types 1 and 4. The result holds for both types of utility functions and in both the cases of paternalistic and welfarist social planner. When some of the households have myopic preferences, the non-monotonic pattern is emphasised: the difference in the replacement rates is bigger.

4. A four-type model, numerical considerations

In this section we generalise the previous model to a four-type economy. The types are indexed as in Table 3. The social planner maximizes a sum of utilities, subject to the revenue constraint (3) and the self-selection constraints given by

$$u(c^i)+\delta^i v(x^i)+\psi(1-y^i)\geq \hat{u}^{ij}(c^j)+\delta^i\hat{v}^{ij}(x^j)+\hat{\psi}^{ij}\left(1-\frac{n^j y^j}{n^i}\right),\ \forall i, j \text{ and } i\neq j.$$

We do not choose the direction of redistribution a priori, i.e. the pattern of the binding self-selection constraints is not restricted a priori, and there appear 12 self-selection constraints in the optimisation problem. Thus, instead of deriving analytical results that are not very tractable in a general case we consider the four-type model only numerically.

Our simulations were also carried out here for CD and CES functions both for welfarist government maximising (3) and paternalistic government maximising (7).[19] First, we assume a uniform distribution of types, i.e. $N^i = 0.25$, implying that there is no correlation between skill level and time preference. Otherwise the parameters are the same as in the two-type case given in Table 1. The results of the numerical solution are given in Tables 7 and 8.

[19] The case with total myopia of types 1 and 2 closely follows the paternalistic case (for details see Tenhunen and Tuomala, 2007). The government with paternalistic views with myopic individuals yields higher marginal tax rates on labour income and savings than in the case without myopia. With myopic individuals in the economy the paternalist government policy increases the future consumption of the low-skilled types and makes the future consumption of the high-skilled types smaller compared to the case without myopia. Also, when interpreting the model as a public pension system, the pension system is more redistributive when myopia is involved.

4.1 Welfarist case

		U	T'	α	x/ny	Consumption dispersion
CD	type 1	-1.48	4.40	0	44.28	means: $\bar{c}$ = 0.91 $\bar{x}$ = 0.67 Lorenz dominance: $GL_c > GL_x$ Gini coefficients: G_c= 0.098 G_x= 0.102
	type 2	-1.59	8.63	-7.08	54.18	
	type 3	-1.11	0	3.20	36.14	
	type 4	-1.19	0	0	43.20	
CES	type 1	-4.73	7.80	0	50.88	means: $\bar{c}$ = 0.72 $\bar{x}$ = 0.62 Lorenz dominance: $GL_c > GL_x$ Gini coefficients: G_c= 0.073 G_x= 0.072
	type 2	-5.10	13.17	-11.38	56.40	
	type 3	-4.22	0	5.14	41.29	
	type 4	-4.52	0	0	45.06	

TABLE 7: Utility levels at the optimum, marginal tax rates, replacement rates and Gini coefficients in the welfarist case. The binding self-selection constraints are ***(3,1), (3,2)*** and ***(4,3)***

In the optimum there are positive marginal labour income tax rates on both low-skilled types. The savings decision of types 1 and 4 is not distorted. Type 2 consumers, who in terms of utilities are the worst-off types, have a positive marginal income tax and a marginal subsidy on savings, whereas type 3, the best-off type in terms of utilities, faces a marginal tax on savings.

Technically speaking, those individuals that are mimicked in the optimum should be taxed. In the two-type case we had a positive marginal tax on both labour income and savings for the mimicked type. With two-dimensional heterogeneity the pattern of binding self-selection constraints can be less straight forward. Similarly to the three-type case, the constraint (4,3) binds here even though type 3 has higher utility than type 4. The direction of other binding constraints (3,1) and (3,2) is something we could expect on the basis of the one-dimensional case with differences only in productivity. We note that the socially most desirable redistribution is from type 3 to type 2 (from the best-off to the worst-off). This binding constraint in turn limits the opportunities to redistribute between other pairs of types. For instance, it would be socially desirable to redistribute from type 4 to type 1 but this is constrained by the fact that type 3 has to be provided with some information "rent" ((4,3) binds). Therefore, if we tax type 4 too heavily he or she may begin to mimic type 3.

4.2. Paternalistic case

		U	T'	α	x/ny	Consumption dispersion
CD	type 1	-1.59	7.95	-10.37	60.08	means: $\bar{c}$ = 0.82 $\bar{x}$ = 0.86 Lorenz dominance: $L_c < L_x$ Gini coefficients: G_c= 0.105 G_x= 0.063
	type 2	-1.65	7.95	-10.37	60.08	
	type 3	-1.14	0	10.35	46.74	
	type 4	-1.16	0	3.23	48.64	
CES	type 1	-4.84	12.81	-17.63	59.58	means: $\bar{c}$ = 0.68 $\bar{x}$ = 0.71 Lorenz dominance: $L_c < L_x$ Gini coefficients: G_c= 0.073 G_x= 0.049
	type 2	-5.15	12.81	-17.63	59.58	
	type 3	-4.23	0	14.13	46.54	
	type 4	-4.50	0	4.49	47.99	

TABLE 8: Utility levels at the optimum, marginal tax rates, replacement rates and Gini coefficients in the paternalistic case. The binding self-selection constraints are ***(3,1), (3,2), (4,3)***

There is bunching of types 1 and 2 in the paternalistic cases, i.e. they always choose a common bundle of labour supply and consumption. This means that the second-order incentive compatibility conditions are not satisfied. This finding gives further validation for the three-type case considered earlier: in the paternalistic case only three different types of consumers can be distinguished.[20]

As in the welfarist case, in terms of utilities type 2 is the worst-off type and type 3 is the best-off type. With the pattern of binding self-selection constraints we could conclude that type 3 mimicking the common choice of types 1 and 2 is prevented by setting a positive marginal income tax on types 1 and 2. This is compensated (at least partly) by a subsidy on savings. The marginal tax on savings for types 3 and 4 is driven by the paternalistic preferences of the social planner.

In the welfarist case the Lorenz curves cross, but for welfare ranking we can use Shorrock's result for the generalised Lorenz curves. For the paternalistic cases we can again apply Atkinson's result. The distribution of income when old is the dominant distribution that has the higher mean. The Gini coefficients for the inequality in the first period consumption and the retirement consumption in Tables 7-8 show that in the second period consumption is less dispersed than in the first period in the paternalistic case. This seems plausible, as government tries to correct the time preference for the second period consumption in the paternalistic cases. In the welfarist case the values of Gini coefficients are practically the same.

4.3. Correlation between skill level and time preference

So far we have assumed a uniform distribution of all types implying that the correlation between skill level and time preference is zero, whereas in the first section the two-type model assumed perfect positive dependence between skill

[20] Blomquist and Christiansen (2004) and Cuff (2000) also apply a three-type version of a Mirrlees (1971) optimal income tax model with heterogeneity in preferences between leisure and consumption.

and time preference. However, this may not be the case; the two characteristics may be imperfectly correlated. At the same time, changing the assumption of the correlation allows us to consider the robustness of our results with respect to the distribution of types. Figures (1) and (2) present the marginal tax rates on savings in the welfarist and paternalist cases with the CES utility function.

From the figures we can see that the results for the marginal tax on savings are fairly robust except for type 3. As the correlation between skill level and time preference increases, or as the fraction of type 2 and 3 decreases, the distortion on type 3's savings decision increases significantly with both types of government preferences. In the welfarist case, type 4 remains undistorted regardless of the structure of the economy. Type 1 is also undistorted except at very high levels of correlation, where his savings become taxed. In the paternalistic case types 1 and 2 remain pooled and the tax on savings for type 4 is almost fixed at all levels of correlation.

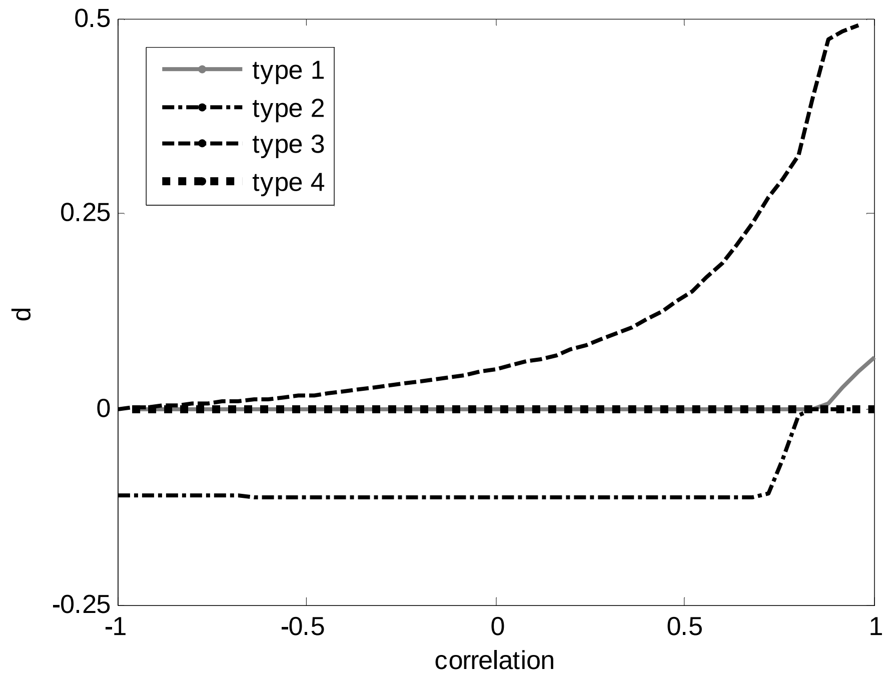

FIGURE 1: Marginal tax on savings in the welfarist case

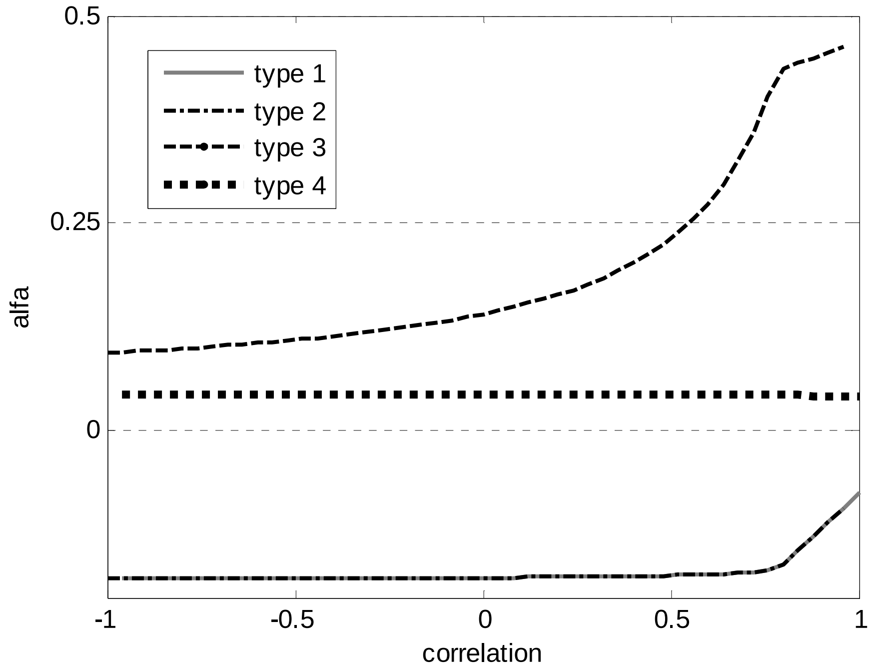

FIGURE 2: Marginal tax on savings in the paternalistic case

5. Conclusions

In this paper we have examined the optimal redistribution policy when society consists of individuals who do not differ only in productivity, but also in time preference or myopia. We extend Diamond's (2003) analysis on the non-linear taxation of savings to the three and four-type models and consider the problem both analytically and numerically. Our results provide a rationale for distortions (upward and downward) in the savings behaviour in a simple two-period model where high-skilled and low-skilled individuals have different non-observable time preferences beyond their earning capacity. When insufficient saving is caused by myopia or low discount factor, our analytical and numerical results support the view that there is a case for a non-linear public pension programme (second-best redistribution) in a world where individuals differ in skills and discount factor or myopia

The numerical results suggest that retirement consumption is less dispersed than first period consumption in the paternalist case, whereas in the welfarist case the ordering is reversed. When the model is interpreted so that there are no private savings, but a public provision of pension in the second period, then in different versions of the model (and with different parameterisations) we find a non-monotonic pattern of the replacement rates. The existence of myopic individuals reinforces the picture. Moreover, with myopic individuals in the economy, in a paternalist case the government policy increases the retirement period consumption of the low-skilled type and decreases the retirement period consumption of the high-skilled types compared to the case without myopic individuals, and thus makes the pension system more redistributive than in a case without myopia.

There are several directions in which we could extend the present model. It would be interesting to explore what kind of picture would emerge if we recognise that the utility from consumption depends on past levels of consumption through habit formation. As in most two-period models of optimum taxation the present paper also assumes that each individual lives for two periods, but is work active, earns income and pays taxes on labour income only in the first period. This is a useful simplification for many purposes, but it means that the model fails to capture the problems arising when an individual is due to pay labour and capital taxes in the same period. For example, the labour supply of the elderly is one of most debated issues in many countries in the context of pension design. Furthermore, if government and individuals have different views on the desirable level of leisure for old people, we may have an additional case for paternalism[21]. To include these concerns we should consider a model in which the individual

[21] Thomas Aronsson suggested this point.

also works and pays labour taxes in the second period when he or she receives the return on capital accumulated in the first period. These questions and possible extensions to the OLG-economy are left for future research.

References

Atkinson, A. B. (1970), 'On the Measurement of Inequality', *Journal of Economic Theory* 2, 244-263.

Atkinson, A. B. and J. E. Stiglitz (1976), 'The Design of Tax Structure: Direct versus Indirect Taxation', *Journal of Public Economics* 1, 55-75.

Berliant, M. and J. O. Ledyard (2005), 'Optimal dynamic nonlinear income taxation with no commitment', working paper.

Blomquist, S. and V. Christiansen (2004), 'Taxation and Heterogeneous Preferences' CESifo Working Paper No.1244.

Cremer, H., P. Pestieau and J-Ch Rochet (2001), 'Direct versus indirect taxation: the design of the tax structure revisited', *International Economic Review* 42, 781-799.

Cremer, H., J.-M. Lozachmeur and P. Pestieau (2004), 'Social Security, retirement age and optimal income taxation', *Journal of Public Economics* 88, 2259-2281.

Cremer, H., P. De Donder, D. Maldonado and P. Pestieau (2006), 'Forced saving, redistribution and non-linear social security scheme', mimeo.

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

Diamond, P. A. (1977), 'A Framework for Social Security Analysis', *Journal of Public Economics* 8, 275-298.

Diamond, P.A. (2003), '*Taxation, Incomplete Markets, and Social Security. The 2000 Munich Lectures*', MIT Press.

Diamond, P. A. and B. Kőszegi (2003), 'Quasi-hyperbolic Discounting and Retirement', *Journal of Public Economics* 87, 1839-1872.

Diamond, P. A. and J. Mirrlees (2000), 'Adjusting One's Standard of Living: Two Period Models' in P. J. Hammond, and G. D. Myles (eds.) '*Incentives, Organization and Public Economics, Papers in Honour of Sir James Mirrlees*', 107-122. Oxford University Press.

Feldstein, M. (1985), 'The Optimal Level of Social Security Benefits', *Quarterly Journal of Economics* 100(2), 300-320.

Gaube, T. (2007). 'Optimum Taxation of Each Year's Income', *Journal of Public Economic Theory* 9, 127-150.

İmrohoroğlu, A., S. İmrohoroğlu and D. H. Joines (2003), 'Time Inconsistent Preferences and Social Security', *Quarterly Journal of Economics* 118(2), 745-784.

Judd, K. L. and C.-L. Su (2006). 'Optimal Income Taxation with Multidimensional Taxpayers Types', Computing in Economics and Finance 2006, 471, Society for Computational Economics.

Kanbur, R. and Tuomala, M. (1994), 'Inherent inequality and the optimal graduation of marginal tax rates', *Scandinavian Journal of Economics* 96 (2), pp. 275-282.

Malinvaud, E. (1972), '*Lectures on Microeconomic Theory'*, North Holland

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

Ramsey, F. P. (1928), 'A Mathematical Theory of Saving', *The Economic Journal* 38, 543-559

Rawls, J. (1971), '*A Theory of Justice'*, Oxford, Oxford University Press.

Roberts, K. (1984), 'The Theoretical Limits to Redistribution', *Review of Economic Studies* 51, 177-195

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', *FinanzArchiv* 50(2), 149-163.

Sen, A.K. (1973), '*On Economic Inequality'*, Clarendon Press, Oxford.

Shorrocks, A. F. (1980), 'The Class of Additively Decomposable Inequality Measures', *Econometrica* 48, 613-626.

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

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

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

Tarkiainen, R and M. Tuomala, (2007), 'On optimal income taxation with heterogeneous work preferences', *International Journal of Economic Theory* 3, 35-46.

Tenhunen, S. and M. Tuomala (2007), 'On Optimal Lifetime Redistribution Policy', Tampere Economic Working Papers, Net Series No. 50.

Tuomala, M. (1990), '*Optimal income tax and redistribution'*, Clarendon Press, Oxford.

Vickrey, W. (1939), 'Averaging of Income for Income Tax Purposes', *Journal of Political Economy* 47, 379-397.

APPENDIX A: First order conditions[22]

The two-type model

The first-order conditions with respect to c^i, x^i and y^i, i=L,H from the Lagrange function given in Eq. (5) are

$$N^L u_c^L - \lambda N^L - \mu \hat{u}_c = 0 \tag{A1}$$

$$N^L \delta^L v_x^L - \lambda r N^L - \mu \delta^H \hat{v}_x = 0 \tag{A2}$$

$$-N^L \psi'^L + \lambda N^L n^L + \mu \hat{\psi}' \frac{n^L}{n^H} = 0 \tag{A3}$$

$$N^H u_c^H - \lambda N^H + \mu u_c^H = 0 \tag{A4}$$

$$N^H \delta^H v_x^H - \lambda r N^H + \mu \delta^H v_x^H = 0 \tag{A5}$$

$$-N^H \psi'^H + \lambda N^H n^H - \mu \psi'^H = 0 \tag{A6}$$

In the paternalistic case the first-order condition with respect to c^i, and y^i, i=L,H, remains unchanged, while (A2) and (A5) are replaced by

$$N^L \delta^g v_x^L - \lambda r N^L - \mu \delta^H \hat{v}_x = 0 \tag{A7}$$

$$N^H \delta^g v_x^H - \lambda r N^H + \mu \delta^H v_x^H = 0 \tag{A8}$$

[22] Note that the functional form for all types is identical. The superscripts used in the derivatives refer to arguments instead of indicating a separate utility function. Similarly, the terms with a 'hat' also refer to arguments, not to different functional forms.

The three-type model

The first order conditions with respect to c^i, x^i and y^i, i=1,3,4 for the problem determined in Eq. (11) are

$$N^1 u_c^1 - \lambda N^1 - \mu^{31} \hat{u}_c^{31} = 0 \tag{A9}$$

$$N^1 \delta^L v_x^1 - \lambda r N^1 - \mu^{31} \delta^L \hat{v}_x^{31} = 0 \tag{A10}$$

$$N^1 \psi'^1 - \lambda N^1 n^1 - \mu^{31} \hat{\psi}'^{31} \frac{n^L}{n^H} = 0 \tag{A11}$$

$$N^3 u_c^3 - \lambda N^3 - \mu^{43} \hat{u}_c^{43} + \mu^{31} u_c^3 = 0 \tag{A12}$$

$$N^3 \delta^L v_x^3 - \lambda r N^3 - \mu^{43} \delta^H \hat{v}_x^{43} + \mu^{31} \delta^L v_x^3 = 0 \tag{A13}$$

$$N^3 \psi'^3 - \lambda N^3 n^H - \mu^{43} \hat{\psi}'^{43} + \mu^{31} \psi'^3 = 0 \tag{A14}$$

$$N^4 u_c^4 - \lambda N^4 + \mu^{43} u_c^4 = 0 \tag{A15}$$

$$N^4 \delta^H v_x^4 - \lambda r N^4 + \mu^{43} \delta^H v_x^4 = 0 \tag{A16}$$

$$N^4 \psi'^4 - \lambda N^4 n^H + \mu^{43} \psi'^4 = 0 \tag{A17}$$

In the paternalistic case (the optimisation problem determined by the Lagrange function given in Eq. (13)) the first-order condition with respect to c^i, and y^i, remains unchanged, while (A10), (A13) and (A16) are replaced by

$$N^1 \delta^g v_x^1 - \lambda r N^1 - \mu^{31} \delta^L \hat{v}_x^{31} = 0 \tag{A18}$$

$$N^3 \delta^g v_x^3 - \lambda r N^3 - \mu^{43} \delta^H \hat{v}_x^{43} + \mu^{31} \delta^L v_x^3 = 0 \tag{A19}$$

$$N^4 \delta^g v_x^4 - \lambda r N^4 + \mu^{43} \delta^H v_x^4 = 0 \tag{A20}$$

Finally, in the myopic case, the optimization problem includes all downwards binding self-selection constraints, so the first order conditions differ from the cases without myopia. The optimum conditions with respect to c^i, x^i and y^i, i=1,3,4 for the problem (15) are

$$N^1 u_c^1 - \lambda N^1 - \mu^{41} \hat{u}_c^{41} - \mu^{31} \hat{u}_c^{31} = 0 \tag{A21}$$

$$N^1 \delta^g v_x^1 - \lambda r N^1 - \mu^{41} \delta^H \hat{u}_c^{41} = 0 \tag{A22}$$

$$N^1 \psi'^1 - \lambda N^1 n^1 - \mu^{41} \hat{\psi}'^{41} \frac{n^L}{n^H} - \mu^{31} \hat{\psi}'^{31} \frac{n^L}{n^H} = 0 \tag{A23}$$

$$N^3 u_c^3 - \lambda N^3 - \mu^{43} \hat{u}_c^{43} + \mu^{31} u_c^3 = 0 \tag{A24}$$

$$N^3 \delta^g v_x^3 - \lambda r N^3 - \mu^{43} \delta^H \hat{v}_x^{43} = 0 \tag{A25}$$

$$N^3 \psi'^3 - \lambda N^3 n^H - \mu^{43} \hat{\psi}'^{43} + \mu^{31} \psi'^3 = 0 \tag{A26}$$

$$N^4 u_c^4 - \lambda N^4 + \mu^{43} u_c^4 + \mu^{41} u_c^4 = 0 \tag{A27}$$

$$N^4 \delta^g v_x^4 - \lambda r N^4 + \mu^{43} \delta^H v_x^4 + \mu^{41} \delta^H v_x^4 = 0 \tag{A28}$$

$$N^4 \psi'^4 - \lambda N^4 n^H + \mu^{43} \psi' + \mu^{41} \psi'^4 = 0 \tag{A29}$$

APPENDIX B: Numerical simulations

Procedure

Numerical simulation was carried out with the Matlab programme. The function used (*fmincon*) solves the optimum of a multivariable function with constraints that may be linear or nonlinear and equality or inequality constraints. It also determines which of the constraints are binding. The same procedure is applied to all cases considered in this paper.

Note that as the optimisation function also allowed slack constraints, we were not restricted to ex ante assumptions on the binding self-selection constraints. Thus, we included all possible constraints in the optimization procedure and simply determined the binding constraints with the help of numerical solutions. In addition to trusting the information given by the program, we also checked that the constraints omitted from the analytical part are indeed slack by calculating the value of the self-selection constraints in question.

Multidimensional screening problems are numerically challenging to solve, as discussed in Judd and Su (2006). We also encountered some difficulties with the solvability of the problem, as the matrix including the constraints with opposite self-selection constraints is very close to being singular. These difficulties were met during the sensitivity analysis; with some combinations of parameters the problem was not solvable, or gave irrational values.[23]

Parameterization

The distribution of the economy was chosen to be uniform merely for the simplicity and comparability for the following cases. The discount factors, δ^L and δ^H are set at 0.6 and 0.8, while δ^g is set equal to 1. These values are in line with e.g. Cremer et al. (2006) doing similar numerical calculations. The myopic cases were modelled by assuming that $\delta^L = 0$.[24] The wage rates reflecting the productivities were chosen to be 2 and 3 respectively. The magnitude of the wage rates was determined by the solvability of the problem; with these wage rates, the labour supply decision, restricted to lie between 0 and 1 were at reasonable levels.

Numerical tables

In addition to those tables presented in the text, we present here some useful numerical values given by the numerical simulations. The first table in each case presents consumptions in the first (c) and the second period (x), the labour supply decision (y) and gross income (ny). The second table provides information on the

[23] A detailed description of cases when the problem is not solvable is provided in Tenhunen and Tuomala (2007).

[24] Note that myopic behaviour does not necessarily imply that δ=0. However, the extreme case with a zero discount factor can be interpreted as a consequence of myopia. In Diamond (2003) myopia is assumed to be with respect to labour supply decisions, but it is not modelled in detail. Our case is just one possibility of the effect of myopia.

binding self-selection constraints. The numerical values for the Lagrange multipliers used in analytical forms of the marginal tax rates are reported here. Figures in the bold face indicate that the constraint in question is binding in the optimum. For the nonbinding constraints the value of the constraint ($U^{ij} – U^{i}$) in the optimum is given in parentheses.[25]

- The two-type case with the welfarist government (Tables B1)

	CD				CES			
	c	x	y	ny	c	x	y	ny
type L	0.8256	0.5022	0.5629	1.1258	0.6667	0.5118	0.4997	0.9994
type H	1.0082	0.8490	0.6639	1.9917	0.7899	0.7248	0.5440	1.6332

	λ	μ^{LH}	μ^{HL}
CD	**1.091**	0 (-4.1493)	**0.050**
CES	**1.872**	0 (-2.8564)	**0.084**

- The two-type case with the paternalistic government (Tables B2)

	CD				CES			
	c	x	y	ny	c	x	y	ny
type L	0.6716	0.7383	0.6221	1.2442	0.5911	0.6287	0.5340	1.0680
type H	0.9668	0.9870	0.6777	2.0331	0.7813	0.7843	0.5489	1.6467

	λ	μ^{LH}	μ^{HL}
CD	**1.221**	0 (*)	**0.090**
CES	**2.084**	0 (-2.9152)	**0.136**

- The three-type case with the welfarist government (Tables B3)

	CD				CES			
	c	x	y	ny	c	x	y	ny
type 1	0.7958	0.5026	0.5662	1.1324	0.6459	0.5133	0.5032	1.0064
type 3	1.1116	0.6839	0.6295	1.8885	0.8301	0.6446	0.5207	1.5621
type 4	1.0162	0.8558	0.6613	1.9839	0.7952	0.7297	0.5409	1.6227

[25] We denote by (*) the case where utility does not get a real value. This is due to the fact that the low-skilled type should supply more labour than the total time endowment. This, of course, is not in the interests of any consumer. Thus, even if the difference cannot be determined accurately, the self-selection constraint in question is not binding.

	λ	μ^{13}	μ^{14}	μ^{31}	μ^{34}	μ^{41}	μ^{43}
CD	**1.0754**	0 (-1.5323)	0 (-3.4165)	**0.0721**	0 (-0.0448)	0 (-0.0616)	**0.0232**
CES	**1.8556**	0 (-1.9739)	0 (-2.6514)	**0.1130**	0 (-0.0362)	0 (-0.0794)	**0.0433**

- The three-type case with the paternalistic government (Tables B4)

	CD				CES			
	c	x	y	ny	c	x	y	ny
type 1	0.6491	0.7541	0.6273	1.2546	0.5759	0.6408	0.5378	1.0756
type 3	0.9962	0.9402	0.6679	2.0037	0.7948	0.7556	0.5411	1.6233
type 4	0.9747	0.9929	0.6751	2.0253	0.7851	0.7872	0.5467	1.6401

	λ	μ^{13}	μ^{14}	μ^{31}	μ^{34}	μ^{41}	μ^{43}
CD	**1.2236**	0 (*)	0 (*)	**0.1029**	0 (-0.0109)	0 (-0.0441)	**0.0482**
CES	**2.0928**	0 (-2.5271)	0 (-2.7585)	**0.1529**	0 (-0.0106)	0 (-0.0474)	**0.0725**

- The three-type case with the paternalistic government and myopic individuals (Tables B5)

	CD				CES			
	c	x	y	ny	c	x	y	ny
type 1	0.6258	0.8168	0.6329	1.2558	0.5607	0.6808	0.5415	1.0830
type 3	1.0419	0.8168	0.6528	1.9584	0.8159	0.6808	0.5289	1.5867
type 4	0.9682	0.9834	0.6779	2.0337	0.7802	0.7828	0.5495	1.6482

	λ	μ^{13}	μ^{14}	μ^{31}	μ^{34}	μ^{41}	μ^{43}
CD	**1.2267**	0 (-2.3626)	0 (*)	**0.0851**	0 (-0.1484)	**0.0309**	**0.0155**
CES	**2.1017**	0 (-2.1014)	0 (-3.0086)	**0.1231**	0 (-0.1531)	**0.0466**	**0.0233**

- The four-type case with the welfarist government (Tables B6)

	CD				CES			
	c	x	y	ny	c	x	y	ny
type 1	0.7998	0.5051	0.5704	1.1408	0.6483	0.5152	0.5063	1.0126
type 2	0.7092	0.6396	0.5902	1.1804	0.6018	0.5828	0.5167	1.0334
type 3	1.1146	0.6814	0.6285	1.8855	0.8318	0.6438	0.5198	1.5594
type 4	1.0172	0.8566	0.6609	1.9827	0.7960	0.7305	0.5404	1.6212

	λ	μ^{12}	μ^{13}	μ^{14}	μ^{21}	μ^{23}	μ^{24}
CD	**1.10**	0 (-0.0256)	0 (-1.5027)	0 (-3.3553)	0 (-0.0216)	0 (-1.4644)	0 (-3.2713)
CES	**1.90**	0 (-0.0275)	0 (-1.9404)	0 (-2.6254)	0 (-0.0175)	0 (-1.8804)	0 (-2.5286)
		μ^{31}	μ^{32}	μ^{34}	μ^{41}	μ^{42}	μ^{43}
CD		**0.03**	**0.06**	0 (-0.0457)	0 (-0.0599)	0 (-0.0127)	**0.03**
CES		**0.05**	**0.08**	0 (-0.0369)	0 (-0.0775)	0 (-0.0325)	**0.05**

- The four-type case with the paternalistic government (Tables B7)

	CD				CES			
	c	x	y	ny	c	x	y	ny
type 1	0.6532	0.7590	0.6317	1.2634	0.5759	0.6408	0.5378	1.0756
type 2	0.6532	0.7590	0.6317	1.2634	0.5759	0.6408	0.5378	1.0756
type 3	1.0034	0.9392	0.6698	2.0094	0.7948	0.7556	0.5411	1.6233
type 4	0.9850	0.9883	0.6773	2.0319	0.7851	0.7872	0.5467	1.6401

	λ	μ^{12}	μ^{13}	μ^{14}	μ^{21}	μ^{23}	μ^{24}
CD	**1.22**	0 (0)	0 (*)	0 (*)	0 (0)	0 (*)	0 (*)
CES	**2.09**	0 (0)	0 (-2.5271)	0 (-2.7585)	0 (0)	0 (-2.4797)	0 (-2.7005)
		μ^{31}	μ^{32}	μ^{34}	μ^{41}	μ^{42}	μ^{43}
CD		**0.05**	**0.05**	0 (-0.0111)	0 (-0.0375)	0 (-0.0375)	**0.05**
CES		**0.08**	**0.08**	0 (-0.0106)	0 (-0.0474)	0 (-0.474)	**0.07**

Sensitivity analysis[26]

The numerical results presented look fairly robust according to the sensitivity analysis done. The first indication of this is that as the simulations are executed with two types of utility functions, Cobb-Douglas and CES, the qualitative results hold.

The two-type case follows the standard results, so that part of the sensitivity analysis is omitted. The sensitivity of the results in the three-type economy for the chosen parameter values of discount factors, group sizes and productivity (wage) differences is considered. With some combinations of the parameters there is a failure to find the optimum, but it is due to technical problems in the optimisation algorithm. The analysis[27] shows that the same qualitative results remain with different parameterization of discount factors, group sizes and productivity rates. With increasing wage inequality the marginal tax rates for both labour income and savings increase, which concurs with the findings in the atemporal continuous Mirrlees model (see Kanbur and Tuomala, 1994). Interestingly, the pattern of the replacement rates does not change. We also consider alternative structures of the economy in Appendix C.

APPENDIX C: Discussion of the groundings for a three-type economy

The three-type economy considered in the text was formed so that all the low productivity types had a low discount factor, whereas some of the high productivity types had a high discount factor and another part had a low discount factor. This, of course, is just one possibility. There are also alternative ways to construct the economy: any group from the structure given in Table 3 might be discarded.

However, there are some aspects that defend three-type analysis. One is the numerical results from the general four-type case. The results suggested that with paternalistic government the low-skilled types 1 and 2 are always pooled. This result seemed to be very insensitive to parameter changes, and a similar result is also found in Cremer et al. (2006). In light of this finding, we have excluded the cases where all high productivity types are pooled, even if it obviously is a possible case.

This finding of pooling of low-productivity types has also been the basis for the parameterization of the structure of the three-type economy. With the low-skilled types pooled, we chose to follow otherwise uniform distribution; half of

[26] A detailed sensitivity analysis for a three-type model is provided in Tenhunen and Tuomala (2007). Calculations of the sensitivity in four-type model are available from the authors upon request.

[27] A more detailed sensitivity analysis can be found in the earlier version of the paper (Tenhunen and Tuomala, 2007).

the economy still consists of low productivity types while half of them have high productivity.

The choice between pooling low productivity types to type 1 or to type 2 still remains to be solved. There is no clear justification for either option, so the choice is made randomly. We have also worked through the alternative case, with no type 1 individuals in the economy and half the economy consisting of type 2 individuals with low productivity and high discount factor. The results follow the same lines as in the case considered in the text with one exception. Now the savings of type 2 are subsidised, even with welfarist government. However, this is not very surprising as the numerical results in the general four-type model suggest that there, too, savings of type 2 were subsidised. Otherwise the results of replacement rates and consumption dispersion remain the same.[28]

[28] The numerical simulations for this alternative case are available from the authors upon request.