



UPPSALA
UNIVERSITET

Uppsala Center for Fiscal Studies

Department of Economics

Working Paper 2011:7

The Marginal Cost of Public Funds is One

Bas Jacobs

Uppsala Center for Fiscal Studies
Department of Economics
Uppsala University
P.O. Box 513
SE-751 20 Uppsala
Sweden
Fax: +46 18 471 14 78

Working paper 2011:7
June 2011

THE MARGINAL COST OF PUBLIC FUNDS IS ONE

BAS JACOBS

The Marginal Cost of Public Funds is One*

Bas Jacobs[†]

November 2, 2010

Abstract

This paper develops a Mirrlees (1971) framework with heterogeneous agents to analyze optimal redistributive taxes, optimal provision of public goods and the marginal cost of public funds (MCF). Standard MCF measures are shown to suffer from three defects: i) The MCF for the (non-individualized) lump-sum tax is generally not equal to one. ii) The MCF for distortionary taxes is not directly related to the marginal excess burden. iii) MCF measures for both lump-sum and distortionary taxes are highly sensitive to the choice of the untaxed numéraire good. These problems are caused by using the *private* rather than the *social* marginal value of private income to calculate the MCF, and disappear by using the *social* marginal value of private income. Moreover, by allowing for redistributive concerns, the marginal excess burden of distortionary taxes equals the marginal distributional gain at the optimal tax system. MCF therefore equals one, both for lump-sum and distortionary taxes, and the modified Samuelson rule should not be corrected for the marginal cost of public funds.

Keywords: marginal cost of funds, marginal excess burden, optimal taxation, optimal redistribution, optimal provision of public goods, Samuelson rule
JEL-codes: H20, H40, H50.

*I thank the Dutch Ministry of Economic Affairs and the Netherlands Organisation for Scientific Research (Vidi-grant No. 452-07-013) for financial support. I am very grateful to Ruud de Mooij, whose contributions to this paper have been invaluable. Parts of this paper have been written when I was a guest researcher at the Center for Fiscal Studies; I thank Uppsala University for its hospitality. Numerous discussions over the years with many people have contributed to the ideas that are developed in this paper. I would like to thank Sören Blomquist, Lans Bovenberg, Vidar Christiansen, Josse Delfgaauw, Casper van Ewijk, Aart Gerritsen, James Heckman, Martin Hellwig, Laurence Jacquet, Etienne Lehmann, Joana Pereira, Rick van der Ploeg, Agnar Sandmo, Dirk Schindler, Paul Tang and Floris Zoutman in particular. Furthermore, I thank seminar participants at the Dutch Ministry of Economic Affairs, CPB Netherlands Bureau for Economic Policy Analysis, NHH-Bergen, Uppsala University, University of St. Gallen, Tinbergen Institute, Erasmus University Rotterdam and Max Planck Institute Bonn for their comments and suggestions. In addition, valuable comments were made by participants at the following congresses: 'Taxation, Public Goods Provision and the Future of the Nordic Welfare Model', Helsinki, June 15, 2009; Congress of the European Economic Association, Barcelona, August 24, 2009; and the 6th CESifo Norwegian-German Seminar on Public Economics, Munich, November 13-14, 2009. Naturally, all errors remain my own.

[†]Bas Jacobs, Erasmus University Rotterdam, Tinbergen Institute, Netspar and CESifo. Address: Erasmus School of Economics, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands. Phone: +31-10-4081481/1491. Fax: +31-10-4089166. E-mail: bjacobs@ese.eur.nl. Homepage: <http://www.xs4all.nl/~jacobs73>.

Pigou (1947): “Where there is indirect damage, it ought to be added to the direct loss of satisfaction involved in the withdrawal of the marginal unit of resources by taxation, before this is balanced against the satisfaction yielded by the marginal expenditure. It follows that, in general, expenditure ought not to be carried out so far as to make the real yield of the marginal unit of resources expended by the government equal to the real yields of the last unit left in the hands of the representative citizen (p. 34).”

Browning (1976): “If the financing of expenditure programs involves a welfare cost, then this cost should be considered part of the opportunity cost of the expenditure programs. ... Thus, an expenditure program will be efficient only if its benefits exceed the direct tax cost by an amount that is at least as large as the additional welfare cost of the funds (p. 283).”

Ballard and Fullerton (1992): “Samuelson’s formula assumes that all of the revenue needed to finance public goods can be raised with lump-sum taxes. Since this is not generally possible, the formula needs to be modified to account for the distortionary effects of the tax system. An appropriate modification is to multiply the cost side of the equation by a term that is commonly called the marginal cost of public funds (MCF) (p. 117).”

1 Introduction

The *marginal cost of public funds* is the ratio of the marginal value of a unit of resources raised by the government and the marginal value of that unit of resources for the private sector. The marginal cost of public funds is therefore a measure indicating the scarcity of public resources. Ever since Pigou (1947), many scholars and policymakers are convinced that the marginal cost of funds must be larger than one, since the government relies on distortionary taxes to finance its outlays. Due to the deadweight losses associated with higher taxes, the private sector needs to sacrifice more welfare (expressed in monetary terms) than the one unit of funds raised through distortionary taxes. The quotes above illustrate well the notion that the marginal cost of funds is larger than one.

If the marginal cost of public funds is indeed larger than one, this has important normative consequences for the determination of optimal public policy in many fields. In early theoretical contributions, Stiglitz and Dasgupta (1971) and Atkinson and Stern (1974) demonstrated that the Samuelson (1954) rule for the optimum provision of public goods needs to be modified to account for tax distortions.¹ The optimal level of public-goods provision should be lower if the marginal cost of public funds is higher. Sandmo (1975) and Bovenberg and de Mooij (1994) demonstrated that the optimal corrective tax is generally set below the Pigouvian tax that perfectly internalizes all externalities, since a high marginal cost of public funds makes the internalization of externalities more expensive. Laffont and Tirole (1993) put the marginal cost

¹Ballard and Fullerton (1992) provide an extensive review of the large literature that subsequently emerged.

of public funds at the center of their theories on optimal procurement and regulation. A larger cost of public funds renders rent extraction more valuable compared to provision of cost-reducing incentives. In the analysis of optimal debt management provided by Barro (1979) and Lucas and Stokey (1983), tax-smoothing ensures that the distortionary cost of taxation is equalized over time. Consequently, debt policy essentially ensures that the marginal cost of public funds is smoothed over time. Some applied policy analysts increase the cost of public projects in social cost-benefit analyses by pre-multiplying the costs of the project with a marginal cost of public funds exceeding one. See, for a recent example, Heckman et al. (2010), who add 50 cents per dollar spent on public programs to correct for the deadweight costs of taxation. Many other examples can be given, but the message is clear: the marginal cost of public funds has a tremendous impact on how governments should evaluate the desirability of public policies.

This paper demonstrates that the scientific debate on the marginal cost of public funds has gone astray in various directions. It reveals that standard marginal cost of public funds measures – commonly derived in linear tax models – suffer from three important defects. Firstly, the marginal cost of funds for lump-sum taxes is not equal to one, whereas there is no theoretical presumption that it should differ from one, given that lump-sum taxes are non-distortionary. Secondly, the marginal cost of public funds for distortionary taxes is generally unrelated to the marginal excess burden of taxation, although this relationship is often suggested (see Pigou, 1947; Harberger, 1964; Browning, 1976). Indeed, one would expect a direct correspondence between the marginal cost of public funds and the deadweight loss of taxation. Thirdly, standard measures for the marginal cost of public funds for both lump-sum and distortionary taxes are disturbingly sensitive to the choice of the untaxed good in the economy. The marginal cost of public funds measures for all tax instruments do not remain the same if the choice of the untaxed numéraire good changes – even though the allocations remain identical under each normalization. These problems are suggestive of a problematic concept of the marginal cost of public funds, which renders the marginal cost of funds measures useless in economic theory and in applied policy analysis.

The analysis of this paper shows that these defects are caused by adopting an economically unappealing measure for the social marginal value of private income when computing the marginal cost of public funds. In particular, to obtain an economically meaningful measure for the marginal cost of funds, the *social* marginal value of a unit of public income should be divided by (the average of) the *social* marginal value of a unit of private income. Instead, the *private* marginal value of private income is commonly used as a measure for the social marginal value of private income.² However, this measure does not correctly measure the social marginal value of private income, since it ignores policy-induced income effects on taxed bases. If the government loses or gains tax revenues due to income effects on taxed bases, these revenue gains or losses create a divergence between *social* marginal value of private income and the *private* marginal value of private income.

This paper follows Diamond (1975) to define the *social* marginal value of private income by the private marginal value of private income *plus* the income effects on taxed bases due to

²It could be multiplied by a welfare weight in order to capture a social preference for redistribution.

transferring resources between the public and private sector. If the marginal cost of public funds is defined as the ratio between the *social* marginal value of public income and the *social* marginal value of private income, this yields a marginal cost of funds measure that is equal to unity for lump-sum taxation. Moreover, a direct correspondence is obtained between the marginal cost of public funds and the marginal excess burden of taxation – as long as distributional concerns are absent. In particular, it is shown that the marginal cost of funds of a distortionary tax then equals the inverse of one minus the excess burden of the tax. Finally, the marginal cost of public funds measures have become insensitive to the normalization of the tax code.

Resolving these theoretical issues is a prerequisite to developing the main contribution of the paper. In particular, the conventional wisdom is challenged that the marginal cost of public funds should be larger than one, due to the distortions caused by taxation. Most of the literature, on which the analysis of the marginal cost of funds has been based, has focused on Ramsey (1927) frameworks with homogeneous agents. Non-individualized lump-sum taxes are ruled out so as to obtain a non-trivial second-best analysis (see, e.g., Browning, 1976, 1987; Wildasin, 1984; Ballard and Fullerton, 1992).³ However, in this context Sandmo (1998) and Werning (2007) note the following:

Sandmo (1998): “The observation that many discussions of the marginal cost of public funds are based on theoretical models where distributional issues have been assumed away is somewhat paradoxical, since the distortionary effects of taxation, on which thinking around the MCF is based, can only be justified from a welfare economics point of view by their positive effects on the distribution of income” (pp. 378–379).

Werning (2007): “Societies may have good reasons for avoiding complete reliance on lump-sum taxes, but none of these are captured by a representative-agent Ramsey framework. Although the first-best allocation is ruled out, an arbitrary second-best problem is set in its place. What confidence can we have that tax recommendations obtained this way accurately evaluate the trade-offs faced by society? If, for unspecified reasons, lump-sum taxes are presumed undesirable, yet still highly desirable within the model, how can we be sure that tax prescriptions derived are not, for the same unspecified reasons, also socially undesirable? (pp. 925–926).”

The literature on optimal income taxation – pioneered by Mirrlees (1971) – provides a solid microeconomic foundation for analyzing tax distortions. Households are heterogeneous with respect to their earning ability (or skill levels). The government aims to maximize a standard Samuelson-Bergson social welfare function, implying a preference for a more equal welfare distribution between households. This can be due to either diminishing marginal utility of income at the household level or a concave transformation of household welfare levels. However,

³Some authors have allowed for heterogeneous agents; see, for example, the contributions by Browning and Johnson (1984) and Allgood and Snow (1998), but these studies focus mainly on the efficiency costs of taxation. Others have explicitly introduced distributional aspects of public goods and taxes (see, e.g., Christiansen, 1981; Wilson, 1991; Boadway and Keen, 1993; Kaplow, 1996; Sandmo, 1998; Slemrod and Yitzhaki, 2001; Dahlby, 2008).

the government cannot achieve any desired distribution of welfare due to informational constraints. Since earning abilities are private information, an individualized lump-sum tax is not incentive-compatible. Therefore, the government needs to rely on a tax on observable income (or consumption), which results in labor supply distortions. Distributional concerns in combination with informational constraints thus provide the fundamental reason why distortionary taxation is optimal in the second-best.

Ramsey (1927) frameworks with representative agents produce misleading results on the marginal cost of public funds by not allowing for distributional concerns to motivate tax distortions. In the Mirrlees (1971) framework, the informational constraint that ability is private information does not exclude the possibility that the government levies non-individualized lump-sum taxes, as these taxes are always incentive-compatible. Hence, non-individualized lump-sum taxes need not be ruled out in an ad hoc fashion so as to obtain a non-trivial second-best policy problem. This paper demonstrates that, at the optimal tax system, the marginal cost of public funds is equalized across all tax instruments. Since the government always has a non-individualized lump-sum tax at its disposal, the marginal cost of public funds for all distortionary taxes should then be equal to the unit marginal cost of public funds of the non-individualized lump-sum tax. The excess burden of income taxes should thus be exactly compensated by the social marginal benefits of redistribution; the tax system would otherwise not have been optimized. Therefore, at the optimal tax system the marginal cost of public funds is equal to one for distortionary taxes as well.

This paper employs a Mirrlees (1971) framework that is extended with public-goods provision to demonstrate that redistribution of income and the provision of public goods should be seen as separate government tasks. The marginal excess burden of distortionary taxation is the *price of equality*, not the price of the government – excluding its redistributive tasks. If the government was not interested in redistribution, then all distortionary tax rates would be zero and all public goods would be financed with non-distortionary non-individualized lump-sum taxes. This finding reveals that redistribution and public-goods provision are not properly disentangled in the Ramsey (1927) approach with a representative agent, where both tasks are (implicitly) lumped together in the government-revenue requirement.

The finding that the marginal cost of public funds equals one has significant normative consequences for the way in which governments should evaluate public policy. Indeed, it is shown to be erroneous to correct the modified Samuelson rule for public-goods provision for the marginal cost of public funds as long as the tax system is optimally set. Consequently, no corrections for deadweight losses of taxation need to be made in social cost-benefit analysis. Extending the line of reasoning of the current paper, Jacobs and de Mooij (2010) demonstrate that it is not desirable to pursue less aggressive environmental policies in the presence of distortionary taxation. Werning (2007) finds that the optimal path of public debt becomes indeterminate and Ricardian equivalence is restored in the presence of distortionary redistributive taxation.

This paper's findings are related to Kaplow (1996), Kaplow (2004), Laroque (2005) and Gauthier and Laroque (2009). These authors explore the conditions under which tax distortions arising from the income tax need not play a role in second-best policy rules under sub-optimal

income taxation. In particular, if the government designs benefit-absorbing changes in the non-linear income tax schedule, it can fully offset the distributional impact of larger public good provision. Moreover, the public good cum tax change does not change labor supply incentives, as long as preferences are identical across households and weakly separable between labor supply and other commodities. Consequently, the benefit-absorbing tax change perfectly imitates a pure benefit tax, and standard first-best policy rules can be applied in the second-best.⁴ In contrast, this paper shows that distortions from labor taxation (i.e. the marginal cost of public funds) should *never* play a role in second-best policy rules, as long as the income tax is optimized. This result holds true without changing the entire tax schedule to neutralize the distributional impact of the public good and without imposing weak separability on preferences. The standard first-best Samuelson rule for the optimal provision of public goods is obtained in second-best with optimal non-linear taxation under homogeneous preferences and weak separability between labor and commodities in the utility function. When the government can only optimize a linear income tax, not only are homogeneous preferences and separability required, but also the absence of income effects in labor supply.

Finally, this paper contributes to the optimal tax literature on optimal (non-linear) taxation and public-goods provision, as in Christiansen (1981), Boadway and Keen (1993), Kaplow (1996) and Sandmo (1998), that assumed identical utility functions for all agents. This paper allows for fully general preference structures, which can be dependent on skill levels.⁵ As the marginal rate of substitution between public and private goods (i.e., the marginal willingness to pay for public goods) is allowed to vary with skill type, it is possible to explore whether public goods should be employed for redistributive purposes alongside the optimal non-linear income tax. It is demonstrated that preference heterogeneity does not fundamentally affect the optimal policy rules under linear income taxation and public-goods provision. However, under non-linear income taxation, public-goods provision is also determined by distributional concerns, which is a novel finding compared to the literature assuming homogeneous preferences. In particular, if the marginal willingness to pay for the public good declines with earning ability, the government will provide more public goods, compared to the first-best Samuelson rule for public-goods provision. The reverse reasoning holds if the willingness to pay for public goods increases with earning ability.⁶

The remainder of this paper is structured as follows. Section 2 introduces the model. Section 3 is devoted to optimal taxation, optimal provision of public goods, and the marginal cost of public funds in heterogeneous-agent settings under linear tax instruments. Section 4 discusses optimal taxation and public-goods provision under non-linear taxation. Section 5 discusses the results and Section 6 concludes.

⁴Kreiner and Verdellin (2010) elaborate on the consequence of non-separable and heterogeneous preferences in similar settings.

⁵This paper is related to Kreiner and Verdellin (2010). These authors explore optimal public good provision with heterogeneous preferences with sub-optimal non-linear income taxation.

⁶These findings extend the earlier results of Saez (2002) on optimal commodity taxation with preference heterogeneity to optimal public-goods provision with preference heterogeneity.

2 Model

A static neoclassical general equilibrium model is considered consisting of households, firms and a government. It is based on earlier contributions by Ramsey (1927), Mirrlees (1971), Diamond and Mirrlees (1971), Sheshinski (1972), Diamond (1975), Christiansen (1981), Boadway and Keen (1993), Kaplow (1996) and Sandmo (1998). Heterogeneous households differ in their earnings capacity and maximize utility by demanding consumption goods and supplying labor effort. Profit-maximizing firms produce private goods and public goods. Both labor and goods markets are perfectly competitive, so that households and firms take wage rates and goods prices as given. The government maximizes social welfare by optimally setting taxes and public-goods provision while taking into account all behavioral responses of households and firms.

Attention is initially focused on the optimal setting of linear tax instruments. Later, also optimal non-linear taxation is explored. Real-world tax systems are neither completely flat, nor completely non-linear. Hence, from a practical point of view it is useful to explore both cases. Moreover, linear tax systems are much less demanding in terms of information. Consequently, it is also of theoretical importance to explore whether the results are robust to reducing the amount of information available to the government; linear tax systems only require verifiability of aggregate labor earnings, whereas non-linear tax systems require verifiability of individual earnings. Since household earnings ability is assumed to be unverifiable to the government, the government cannot employ individualized lump-sum taxes or transfers.

2.1 Households

There is a mass N of households, which may differ by a single-dimensional parameter $n \in \mathcal{N} = [\underline{n}, \bar{n}]$, where the upper bound \bar{n} could be infinite. n denotes the household's earning ability ('skill level') and represents the endowment of efficiency units of labor of each household. The density of household types with index n is denoted by $f(n)$ and the cumulative distribution function by $F(n)$. Note that $N \int_{\mathcal{N}} f(n) dn = N$.

Each household n derives utility $u(n)$ from consumption $c(n)$, and pure public goods G . It derives disutility from labor $l(n)$. Each household has an endowment of time, which is allocated between leisure and working. Consumption and leisure are both assumed to be normal goods. Preferences can be heterogeneous, as the utility function may depend on the skill level n . This generalizes the previously mentioned papers on optimal taxation and public-goods provision that assume homogeneous preferences.

The utility function satisfies the Inada conditions, is strictly concave, and is twice continuously differentiable:

$$u(n) \equiv u(c(n), l(n), G; n), \quad u_c, -u_l, u_G > 0, \quad u_{cc}, u_{ll}, u_{GG} < 0, \quad \forall n. \quad (1)$$

Subscripts denote partial derivatives. The Spence-Mirrlees conditions are imposed on the utility function so as to guarantee the implementability of the non-linear policies derived in the next section (see also Lemma 2 below). No other restrictions are imposed on the derivatives of the

utility function.

All skill types are assumed to be perfect substitutes in production; hence, the skill price per efficiency unit of labor is equal across households and denoted by w .⁷ The government employs a linear tax schedule consisting of a linear tax rate t on gross labor earnings $z(n) \equiv wnl(n)$, a linear tax rate τ on consumption goods $c(n)$, and a non-individualized lump-sum transfer T . The informational assumptions for linear taxes are that the government should be able to verify aggregate labor income or consumption. Note that in Mirrlees (1971)-analyses non-individualized lump-sum transfers are always part of the instrument set of the government, since the government can always provide each household with an equal amount of resources.

The household budget constraint states that expenditures on consumption are equal to net labor earnings plus non-individualized lump-sum transfers:

$$(1 + \tau)c(n) = (1 - t)z(n) + T, \quad \forall n. \quad (2)$$

One tax instrument is redundant, since the consumption tax is equivalent to the income tax. Thus, without loss of generality, one tax instrument can always be normalized to zero.

The household maximizes utility (1) subject to its budget constraint (2). This yields the standard first-order condition for labor supply:⁸

$$\frac{-u_l(c(n), l(n), G; n)}{u_c(c(n), l(n), G; n)} = \frac{(1 - t)wn}{(1 + \tau)}, \quad \forall n. \quad (3)$$

Taxation is distortionary as it drives a wedge between the marginal social benefits (wn) and the marginal private benefits ($\frac{(1-t)wn}{(1+\tau)}$) of an increase in labor effort. Higher levels of taxation (income or consumption taxes) reduce the price of leisure in terms of consumption goods and households substitute leisure for consumption.

Indirect utility of household n can be written as: $v(n) \equiv v(t, \tau, T, G; n) \equiv u(\hat{c}(n), \hat{l}(n), G; n)$, where hats denote the optimized values for consumption and labor supply. Straightforward application of Roy's identity yields the following properties of $v(\cdot)$: $\frac{\partial v(n)}{\partial t} = -\lambda(n)z(n)$, $\frac{\partial v(n)}{\partial \tau} = -\lambda(n)c(n)$, $\frac{\partial v(n)}{\partial T} = \lambda(n)$ and $\frac{\partial v(n)}{\partial G} = \lambda(n)\frac{u_G}{u_c}$, where $\lambda(n)$ is the private marginal utility of income of household n .

Some additional notation is introduced to express the optimal policy rules in terms of well-known elasticity concepts. In particular, the compensated, uncompensated and income elasticities of labor supply and consumption demand with respect to the tax rates and the public good are denoted by: $\varepsilon_{lt}^u \equiv \frac{\partial l^u(n)}{\partial t} \frac{1-t}{l(n)}$, $\varepsilon_{lt}^c \equiv \frac{\partial l^c(n)}{\partial t} \frac{1-t}{l(n)} < 0$, $\varepsilon_{lG}^u \equiv \frac{\partial l^u(n)}{\partial G} \frac{G}{l(n)}$, $\varepsilon_{lG}^c \equiv \frac{\partial l^c(n)}{\partial G} \frac{G}{l(n)}$, $\varepsilon_{lT} \equiv (1 - t)wn \frac{\partial l(n)}{\partial T} < 0$, $\varepsilon_{c\tau}^u \equiv \frac{\partial c^u(n)}{\partial \tau} \frac{1+\tau}{c(n)}$, $\varepsilon_{c\tau}^c \equiv \frac{\partial c^c(n)}{\partial \tau} \frac{1+\tau}{c(n)}$, $\varepsilon_{cG}^u \equiv \frac{\partial c^u(n)}{\partial G} \frac{G}{c(n)}$, $\varepsilon_{cG}^c \equiv \frac{\partial c^c(n)}{\partial G} \frac{G}{c(n)}$,

⁷The results in this paper are conditional on the assumption that all labor types are perfect substitutes in production. Pirttilä and Tuomala (2001) demonstrate that the Samuelson rule is modified (further) to capture the effect of policy-induced general equilibrium effects on the wage structure. No general statements can be made with regard to the direction in which the optimal taxes or the provision of public goods should be adjusted. This depends on cross-substitution patterns in the production function(s).

⁸Under linear instruments, first-order conditions are both necessary and sufficient to describe the maximum of utility given the restrictions on the derivatives of the utility function. Under non-linear policies, monotonicity and Spence-Mirrlees conditions are required for a maximum; see also section 4.3.

and $\varepsilon_{cT} \equiv (1 + \tau) \frac{\partial c(n)}{\partial T} > 0$, where the superscript u (c) denotes an uncompensated (compensated) change. In addition, the Slutsky equations will be employed to simplify the optimal tax expressions:

$$\frac{\partial l^u(n)}{\partial t} = \frac{\partial l^c(n)}{\partial t} - z(n) \frac{\partial l(n)}{\partial T}, \quad (4)$$

$$\frac{\partial c^u(n)}{\partial \tau} = \frac{\partial c^c(n)}{\partial \tau} - c(n) \frac{\partial c(n)}{\partial T}, \quad (5)$$

$$\frac{\partial l^u(n)}{\partial G} = \frac{\partial l^c(n)}{\partial G} + \frac{u_G}{\lambda(n)} \frac{\partial l(n)}{\partial T}, \quad (6)$$

$$\frac{\partial c^u(n)}{\partial G} = \frac{\partial c^c(n)}{\partial G} + \frac{u_G}{\lambda(n)} \frac{\partial c(n)}{\partial T}, \quad (7)$$

Rewriting the Slutsky equations yields $\varepsilon_{lt}^u = \varepsilon_{lt}^c - \varepsilon_{lT}$, $\varepsilon_{c\tau}^u = \varepsilon_{c\tau}^c - \varepsilon_{cT}$, $\varepsilon_{lG}^c = \varepsilon_{lG}^u - \frac{\omega_G}{\omega_l} \varepsilon_{lT}$ and $\varepsilon_{lG}^c = \varepsilon_{lG}^u - \frac{\omega_G}{\omega_c} \varepsilon_{lT}$, where $\omega_G \equiv \frac{u_G G}{u}$, $\omega_l \equiv \frac{-u_l l(n)}{u}$ and $\omega_c \equiv \frac{u_c c(n)}{u}$ denote the shares of public goods, labor and consumption in utility. In the remainder of the paper, a bar is used to indicate an income-weighted elasticity, e.g., $\bar{\varepsilon}_{lt}^c \equiv \left[\int_{\mathcal{N}} \varepsilon_{lt}^c z(n) dF(N) \right] \left[\int_{\mathcal{N}} z(n) dF(n) \right]^{-1}$.

2.2 Firms

A representative firm produces private and public goods, taking prices for inputs and outputs as given.⁹ C denotes aggregate output of consumption goods. The price of the consumption good is normalized to one without loss of generality, whereas the price of the public good is denoted by p . The firm operates linear production technologies for both goods, with labor being the only input in production:¹⁰

$$C = L_C, \quad G = AL_G, \quad A > 0. \quad (8)$$

L_C and L_G denote labor input in consumption and public-goods production, respectively. Labor efficiency in consumption-goods production is harmlessly normalized to one. A therefore denotes the relative labor efficiency of public-goods production. The zero-profit condition ensures that the equilibrium-wage rate equals unity ($w = 1$). In addition, the relative price of public goods is constant and equals the marginal rate of transformation between public and private goods ($p = 1/A$). The relative price of public goods p is larger (smaller) than one if public-goods production is less (more) efficient than private-goods production: $A < 1$ ($A > 1$).

2.3 Government

The government's objectives are represented by a Samuelson-Bergson social welfare function, which is a concave sum of household utilities:

$$\mathcal{W} \equiv N \int_{\mathcal{N}} \Psi(u(n)) dF(n), \quad \Psi'(u(n)) > 0, \quad \Psi''(u(n)) \leq 0. \quad (9)$$

⁹Alternatively, we could assume that there is a separate public firm producing the public good, but this does not affect any of the results Diamond and Mirrlees (1971).

¹⁰In the absence of constant returns to scale, a perfect profit tax is required to preserve the results of this paper; see also Diamond and Mirrlees (1971).

Maximizing social welfare under a Utilitarian criterion ($\Psi' = 1$) implies a social preference for redistribution if the private marginal utility of income $\lambda(n)$ declines with skill n at the household level. However, due to the concavity of $\Psi(\cdot)$ the government may exhibit a stronger preference for redistribution than households do.

The government's budget constraint states that total tax revenues equal spending on transfers and public goods:

$$N \int_{\mathcal{N}} (tz(n) + \tau c(n)) dF(n) = NT + pG. \quad (10)$$

2.4 General equilibrium

Clearing of both labor and goods markets requires that total labor demand $L_C + L_G$ equals total labor supply, and that total supply of private consumption goods equals total demand for private consumption goods:

$$L_C + L_G = N \int_{\mathcal{N}} nl(n) dF(n), \quad (11)$$

$$C = N \int_{\mathcal{N}} c(n) dF(n). \quad (12)$$

3 Optimal linear taxation and public-goods provision

The government maximizes social welfare subject to its budget constraint by choosing the non-individualized lump-sum transfer, the tax rate on income or consumption and the level of public-goods provision. Optimal policies are derived with both tax normalizations (i.e., either consumption or income remains untaxed). The social marginal value of one unit of public resources is denoted by η .

The Lagrangian for maximizing social welfare is given by

$$\mathcal{L} \equiv N \int_{\mathcal{N}} \Psi(v(n)) dF(n) + \eta \left(N \int_{\mathcal{N}} (tnl(n) + \tau c(n)) dF(n) - NT - pG \right), \quad (13)$$

The first-order conditions for a maximum are

$$\frac{\partial \mathcal{L}}{\partial T} = N \int_{\mathcal{N}} \left(\Psi'(v(n)) \lambda(n) - \eta + \eta tn \frac{\partial l(n)}{\partial T} + \eta \tau \frac{\partial c(n)}{\partial T} \right) dF(n) = 0, \quad (14)$$

$$\frac{\partial \mathcal{L}}{\partial t} = N \int_{\mathcal{N}} \left(-\Psi'(v(n)) \lambda(n) nl(n) + \eta ml(n) + \eta tn \frac{\partial l(n)}{\partial t} + \eta \tau \frac{\partial c(n)}{\partial t} \right) dF(n) = 0, \quad (15)$$

$$\frac{\partial \mathcal{L}}{\partial \tau} = N \int_{\mathcal{N}} \left(-\Psi'(v(n)) \lambda(n) c(n) + \eta c(n) + \eta tn \frac{\partial l(n)}{\partial \tau} + \eta \tau \frac{\partial c(n)}{\partial \tau} \right) dF(n) = 0, \quad (16)$$

$$\frac{\partial \mathcal{L}}{\partial G} = N \int_{\mathcal{N}} \left(\Psi'(v(n)) u_G + \eta tn \frac{\partial l(n)}{\partial G} + \eta \tau \frac{\partial c(n)}{\partial G} \right) dF(n) - p\eta = 0, \quad (17)$$

where Roy's lemma has been used in each first-order condition, and $w = 1$ is imposed.¹¹

¹¹Diamond and Mirrlees (1971) point out that the first-order conditions are necessary, but not generally sufficient, to characterize the optimum. In the remainder it's assumed that the second-order conditions for

3.1 Marginal cost of public funds – The standard approach

This section first derives optimal policies using the standard concept for the marginal cost of funds. To that end, the following definitions are introduced: *marginal social value of private income*, the *marginal cost of public funds*, and the Feldstein (1972) *distributional characteristics* of the tax bases and public goods.

Definition 1 Standard definition of the social marginal value of private income – *The social marginal value of one unit of private income accruing to household n equals*

$$\alpha'(n) \equiv \Psi'(u(n)) \lambda(n), \quad \forall n. \quad (18)$$

According to the traditional definition, the social marginal value of private income is captured by the mechanical increase in social welfare of transferring one unit of resources to household n . Private welfare rises with $\lambda(n)$, and this increase in private welfare is valued by the government by $\Psi'(u(n))$.

Definition 2 Marginal cost of public funds – *The marginal cost of public funds based on the traditional measure of the social marginal value of income is given by*

$$MCF' \equiv \frac{\eta}{\int_{\mathcal{N}} \alpha'(n) dF(n)}. \quad (19)$$

The traditional marginal cost of public funds MCF' is the ratio of the social marginal value of public resources η and the *average* of the social marginal value of private income $\alpha'(n)$.

Definition 3 Distributional characteristics of the tax bases – *The distributional characteristics ξ'_y of tax bases $y(n) = \{z(n), c(n)\}$ are given by*

$$\xi'_y \equiv 1 - \frac{\int_{\mathcal{N}} \alpha'(n) y(n) dF(n)}{\int_{\mathcal{N}} \alpha'(n) dF(n) \int_{\mathcal{N}} y(n) dF(n)} > 0. \quad (20)$$

ξ'_y represents the gain in social welfare (expressed in monetary equivalents and then divided by the taxed base) of redistributing a marginal unit of resources through base $y(n) = \{z(n), c(n)\}$. The distributional characteristics of the tax bases are positive, since the covariance between tax base $y(n)$ and the social welfare weights $\alpha'(n)$ are negative. Households with higher incomes or consumption levels feature lower welfare weights because the social marginal utility of income is diminishing due to diminishing private marginal utility of income and/or concavity of the social welfare function. The positive distributional characteristic ξ'_y therefore implies that redistribution through taxing income or consumption yields distributional benefits. A stronger social desire for redistribution of welfare or greater inequality in the skill distribution raises the distributional characteristic. Since ξ'_y is a positive normalized covariance, it ranges between zero and one. $\xi'_y = 0$ is obtained either if the government is not interested in redistribution and attaches the same social welfare weights $\alpha'(n)$ to all households or if the base $y(n)$ is the same

maximization of social welfare are always satisfied.

for all n so that there is no inequality. $\xi'_y = 1$ if the government pursues a Rawlsian social welfare function.

Definition 4 Distributional characteristic of the public good – *The distributional characteristic of the public good is*

$$\xi'_G \equiv 1 - \frac{\int_{\mathcal{N}} \alpha'(n) \frac{u_G(\cdot)}{u_c(\cdot)} dF(n)}{\int_{\mathcal{N}} \alpha'(n) dF(n) \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n)}. \quad (21)$$

The distributional characteristic for the public good ξ'_G is the negative normalized covariance between the social marginal valuation of income $\alpha'(n)$ and the marginal willingness to pay for the public good $\frac{u_G}{u_c}$. $\xi'_G > 0$ if the public good mainly benefits the rich, who feature the lowest social welfare weights $\alpha'(n)$, and vice versa. $\xi'_G = 0$ if the government is not interested in redistribution and attaches the same welfare weights $\alpha'(n)$ to all households or if all households benefit equally from the public good, so that $\frac{u_G}{u_c}$ is equal for all n .

The next lemma derives the marginal excess burden of the income (or consumption) tax. The excess burden measures the reduction in social welfare, measured in monetary units, expressed as a fraction of the tax base, of raising the distortionary income or consumption tax.

Lemma 1 Marginal excess burden – *The marginal excess burdens for the income tax and the consumption tax are given by*

$$MEB_t = -\frac{t}{1-t} \bar{\varepsilon}_{lt}^c, \quad MEB_\tau = -\frac{\tau}{1+\tau} \bar{\varepsilon}_{c\tau}^c. \quad (22)$$

Proof. The welfare effect of a rise in the tax rate is evaluated, while public-goods provision remains constant. All revenue raised from household n is transferred back to household n through a household-specific lump-sum transfer $s(n)$.¹² All welfare changes are converted in monetary equivalents and then summed over all households without discounting for utility differences through the social welfare function $\Psi(\cdot)$. This ensures that only the efficiency losses of the distorting tax are valued and not its distributional gains. To determine the excess burden of the income tax, assume that $\tau = 0$. The change in taxes dt and lump-sum income $ds(n)$ for each household n satisfies $ds(n) = \left(nl(n) + tn \frac{\partial l(n)}{\partial t} \right) dt$, which leaves the public budget unaffected. The change in utility of household n is given by $dv(n) = \lambda(n) ds(n) - \lambda(n) nl(n) dt$, where Roy's identity is employed. Substituting for $ds(n)$ yields: $\frac{dv(n)/dt}{\lambda(n)} = \frac{t}{1-t} \bar{\varepsilon}_{lt}^c nl(n)$. Since each household gets compensated for the tax change, the behavioral response is measured by the compensated labor supply elasticity. The change in social welfare (expressed in monetary units) is the sum of changes in the welfare of households (in monetary units): $N \int_{\mathcal{N}} \frac{dv(n)/dt}{\lambda(n)} dF(n)$. Hence, the marginal excess burden of the tax (expressed in monetary equivalents, as a fraction of the tax base) is given by:

$$MEB_t \equiv -\frac{\int_{\mathcal{N}} \frac{dv(n)/dt}{\lambda(n)} dF(n)}{\int_{\mathcal{N}} nl(n) dF(n)} = -\frac{t}{1-t} \bar{\varepsilon}_{lt}^c. \quad (23)$$

¹²Of course this instrument does not exist, since it boils down to an individualized lump-sum tax. However, this thought-exercise allows us to calculate the excess burden of the tax.

Recall that the bar indicates an income-weighted elasticity. Using similar steps and setting $t = 0$ gives the excess burden of the consumption tax:

$$MEB_\tau \equiv -\frac{\int_{\mathcal{N}} \frac{dv(n)/d\tau}{\lambda(n)} dF(n)}{\int_{\mathcal{N}} c(n) dF(n)} = -\frac{\tau}{1+\tau} \bar{\varepsilon}_{c\tau}^c. \quad (24)$$

■

Proposition 1 replicates Sandmo (1998) by characterizing optimal tax policies and public-goods provision under the standard concept for the marginal cost of funds, and assuming that the consumption tax is normalized to zero. Sandmo (1998) adds distributional concerns to the analysis of Atkinson and Stern (1974). The first-order conditions for T (14) and t (15) both yield a measure for the marginal cost of public funds. Therefore, MCF'_T and MCF'_t are used to denote the marginal cost of public funds of the lump-sum tax and the tax rate, respectively; see also Sandmo (1998).

Proposition 1 *With the standard MCF definition, and with the consumption tax normalized to zero, the optimal rules for public-goods provision and the linear income tax are characterized by*

$$(1 - \xi'_G) N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = (1 - \gamma_t \bar{\varepsilon}_{tG}^u) \cdot MCF'_t \cdot p, \quad (25)$$

$$MCF'_T = \frac{1}{1 - \frac{t}{1-t} \bar{\varepsilon}_{tT}} < 1, \quad \bar{\varepsilon}_{tT} \equiv \int_{\mathcal{N}} \varepsilon_{tT} dF(n) < 0, \quad (26)$$

$$MCF'_t = \frac{1 - \xi'_z}{1 + \frac{t}{1-t} \bar{\varepsilon}_{tt}^u} = \frac{1 - \xi'_z}{1 - MEB_t - \frac{t}{1-t} \bar{\varepsilon}_{tT}}, \quad (27)$$

$$MCF' = MCF'_T = MCF'_t < 1. \quad (28)$$

Proof. Equation (14) can be simplified by setting $\tau = 0$ and substituting equation (19) to find equation (26). Equation (15) is simplified by using equation (20) and setting $\tau = 0$ to find the first part of equation (27). The second part of (27) follows upon substitution of the Slutsky equation $\bar{\varepsilon}_{tt}^u = \bar{\varepsilon}_{tt}^c - \bar{\varepsilon}_{tT}$ and using equation (22). Equation (17) is simplified by using equation (21), setting $\tau = 0$, and using $\gamma_t \equiv Nt \int_{\mathcal{N}} nl(n) dF(n) / pG$ to find equation (25). ■

Equation (25) is the modified Samuelson rule for the optimal provision of public goods. The modified Samuelson rule equates the sum of the marginal social benefits with the marginal social costs of providing the public good. The benefits $N \int_{\mathcal{N}} \frac{u_G}{u_c} dF(n)$ are deflated by the distributional characteristic of the public good ξ'_G . If poor households value the public good more (less) than rich households do, then $\xi'_G < 0$ ($\xi'_G > 0$). In that case, the level of public-goods provision increases (decreases) – ceteris paribus. The right-hand side gives the marginal cost of providing the public good. p equals the marginal rate of transformation between the public good and the reference private good. The cost side of the Samuelson rule is corrected for tax distortions via the marginal cost of funds MCF' . A higher cost of public funds lowers the level of public-goods provision – ceteris paribus. Providing the public good may exacerbate (reduce) pre-existing labor-tax distortions if the public good reduces (stimulates) labor supply. This correction is governed by the uncompensated cross-elasticity of labor with the public good (if $\bar{\varepsilon}_{tG}^u \neq 0$); see

also Atkinson and Stern (1974). $\gamma_t \equiv Nt \int_{\mathcal{N}} z(n) dF(n) / pG$ denotes the share of public goods that is financed with distortionary taxes. If $\gamma_t = 1$, then public goods are financed exclusively with distortionary taxes. If the public good and labor supply are complementary ($\bar{\varepsilon}_{lG}^u > 0$), the provision of the public good boosts labor supply and this raises tax revenue in the presence of a positive labor tax ($\gamma_t > 0$). These positive revenue effects decrease the cost of public-goods provision. If the public good reduces labor supply ($\bar{\varepsilon}_{lG}^u < 0$), the cost of the public good rises, since the public good reduces labor supply, which erodes the labor tax base.

Equation (26) gives the marginal cost of funds for the lump-sum tax. If there is a positive income tax ($t > 0$), the marginal cost of funds for the lump-sum tax is always smaller than one ($MCF'_T < 1$), as long as leisure is a normal good ($\tilde{\varepsilon}_{lT} < 0$). Intuitively, transferring an extra unit of funds from households to the government via a larger lump-sum tax generates an income effect in labor supply, and this raises tax revenues from the income tax if it is positive ($t > 0$). However, one would expect the lump-sum tax to have a marginal cost of funds equal to one. The result that the lump-sum tax does not have a marginal cost of funds of unity is the first problem that is identified in the literature.

Equation (27) gives the marginal cost of funds for the tax rate (MCF'_t). MCF'_t depends on the income-weighted *uncompensated* tax elasticity of labor supply $\bar{\varepsilon}_{lt}^u$. The sign of this elasticity is ambiguous due to offsetting income and substitution effects.¹³ Clearly, by allowing for distributional concerns ($\xi'_z > 0$), distortionary taxes have not only costs, but also distributional benefits. These distributional gains lower the marginal cost of funds for the distortionary income tax.

In Ramsey (1927) frameworks the non-individualized lump-sum tax is ruled out and distributional concerns are absent. Hence, MCF'_T is not applicable and $\xi'_z = 0$ in MCF'_t ; see also Atkinson and Stern (1974). $MCF'_t > 1$ is then obtained if the labor supply curve is upward-sloping in wages ($\bar{\varepsilon}_{lt}^u < 0$) and $MCF'_t < 1$ if there is a backward-bending labor supply curve ($\bar{\varepsilon}_{lt}^u > 0$). In the absence of distributional concerns (i.e., $\xi'_z = 0$), equation (27) demonstrates that it is impossible to relate the marginal cost of public funds of the distortionary income tax to the marginal excess burden of the income tax, as suggested by the analyses in Pigou (1947), Harberger (1964) and Browning (1976). The often-used approximation that the marginal cost of funds equals one plus the marginal excess burden of the tax will only apply as long as income effects are absent in labor supply ($\bar{\varepsilon}_{lT} = 0$), since then $\bar{\varepsilon}_{lt}^c = \bar{\varepsilon}_{lt}^u$ and $MCF'_t = [1 - MEB_t]^{-1} \approx 1 + MEB_t$ at low levels of t . However, assuming away income effects is theoretically not very appealing and is not empirically warranted; see also Blundell and MaCurdy (1999). Thus, in the absence of distributional benefits of taxes, the marginal cost of funds of the distortionary tax is not directly related to the excess burden of the tax instrument. This is the second problem in the literature on the marginal cost of public funds.

The results that (in the absence of redistributive concerns) the MCF'_t can be smaller than one and that it cannot be directly related to the marginal excess burden of the tax led to a large literature trying to resolve these counterintuitive findings; see for example Triest (1990), Ballard

¹³Most empirical studies conducted in recent decades suggest a positive value for the uncompensated wage elasticity of labor supply, implying that MCF'_t will exceed one; see, for example, Blundell and MaCurdy (1999).

and Fullerton (1992) and Dahlby (2008).

Equation (28) states that, in a second-best optimum, the marginal cost of funds for all tax instruments should be equalized (i.e., $MCF' = MCF'_T = MCF'_t$). Consequently, the marginal cost of funds at the optimal tax system is smaller than one. Note that this result holds as long as there is a positive labor tax, irrespective of whether there are distributional concerns or not. This finding led Sandmo (1998) to conclude that in a second-best setting with redistributive concerns, the marginal cost of public funds is lower than one, rather than higher than one.

The next proposition demonstrates that the standard definition for the marginal cost of funds is highly sensitive to the normalization of the tax system. In particular, it gives the optimality conditions for the tax system and public-goods provision when the income tax is normalized to zero and consumption is taxed, instead. In this case, MCF'_T and MCF'_τ are used to denote the marginal cost of public funds of the lump-sum tax and the consumption tax rate that follow from first-order conditions (14) and (16), respectively.

Proposition 2 *With the standard MCF definition, and with the income tax normalized to zero, the optimal rules for public-goods provision and the linear consumption tax are characterized by*

$$(1 - \xi'_G) N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = (1 - \gamma_\tau \bar{\varepsilon}_{cG}^u) \cdot MCF' \cdot p, \quad (29)$$

$$MCF'_T = \frac{1}{1 - \frac{\tau}{1+\tau} \tilde{\varepsilon}_{cT}} > 1, \quad \tilde{\varepsilon}_{cT} \equiv \int_{\mathcal{N}} \varepsilon_{cT} dF(n) > 0, \quad (30)$$

$$MCF'_\tau = \frac{1 - \xi'_c}{1 + \frac{\tau}{1+\tau} \bar{\varepsilon}_{c\tau}^u} = \frac{1 - \xi'_c}{1 - MEB_\tau - \frac{\tau}{1+\tau} \bar{\varepsilon}_{cT}}, \quad (31)$$

$$MCF' = MCF'_T = MCF'_\tau > 1. \quad (32)$$

Proof. Equation (14) can be simplified by setting $t = 0$ and substituting equation (19) to find equation (30). Equation (15) is simplified by using equation (20), and setting $t = 0$ to find the first part of equation (31). The second part of (31) follows upon substitution of the Slutsky equation $\bar{\varepsilon}_{c\tau}^u = \bar{\varepsilon}_{c\tau}^c - \bar{\varepsilon}_{cT}$ and using equation (22). Equation (17) is simplified by using equation (21), setting $t = 0$, and using $\gamma_\tau \equiv N\tau \int_{\mathcal{N}} c(n) dF(n) / pG$ to find equation (29). ■

The interpretations of equations (29) and (31) are identical to those of (25) and (27) that appear below proposition 1. Hence, these will not be repeated.

Equation (30) reveals that the marginal cost of funds for the lump-sum transfer is now always higher than (or equal to) one, rather than always smaller than (or equal to) one under income taxation. The reason is that the income elasticity of consumption is positive (if consumption is a normal good): $\tilde{\varepsilon}_{cT} > 0$. Hence, MCF'_T switches from a number below one, to a number above one, when a different normalization of the tax system is adopted.

In Ramsey (1927) frameworks with a representative agent ($\xi'_c = 0$) and no lump-sum taxes, equation (31) gives the marginal cost of funds. It reveals that $MCF'_\tau > 1$, since substitution effects and income effects in consumption demand are reinforcing rather than offsetting. Consequently, the marginal cost of funds for the distortionary tax under income normalization is always larger than one, whereas it could be smaller than one under consumption normaliza-

tion. Hence, the marginal cost of funds for the distortionary tax instrument is larger under consumption taxation than under income taxation.¹⁴

In the presence of distributional concerns ($\xi'_c > 0$), equation (32) shows that at the optimal tax system the marginal cost of funds for all tax instruments are equalized. As a result, MCF' is larger than one when consumption is taxed, and MCF' is smaller than one if income is taxed.

A different normalization of the tax system therefore produces completely different marginal cost of funds measures for both lump-sum and distortionary taxes. However, the normalization of the tax system does not influence the optimal second-best allocation. This is the third problem with the standard marginal cost of funds measures.

The normalization explains the findings of Wilson (1991) and Sandmo (1998). They both suggest that distributional concerns explain why the marginal cost of public funds is smaller than one. However, their findings stem from the normalization of the tax system, and would be reversed if consumption would be taken as the tax base, as proposition 4 demonstrates.

To summarize, the standard MCF measure suffers from three defects. Firstly, the marginal cost of funds for lump-sum taxes are not equal to one – irrespective of the normalization of the tax system. Secondly, in the absence of distributional concerns the marginal cost of funds for a distortionary tax instrument cannot be related to the excess burden of the tax instrument – irrespective of the normalization of the tax system. Hence, the MCF measure does not properly capture the welfare costs of taxation. Thirdly, the MCF measures for both distortionary and lump-sum taxes are shown to be critically sensitive to the normalization of the tax system. All these problems are suggestive of a conceptually flawed MCF measure. From a practical point of view, these problems render the standard MCF measures useless in any applied analysis.

3.2 Marginal cost of public funds – The Diamond approach

This section argues that the three problems with the standard MCF measures can be resolved by appropriately specifying the social value of private income, which is used in the definition of the marginal cost of funds. In particular, it is argued that the *social* value of private income should also include the income effects on the taxed bases, as in Diamond (1975). If the government transfers one unit of funds to the private sector, it should not only take into account the direct utility gains to households – as the standard measure does – but also include the indirect effect of a higher transfer on other tax revenues as social losses (under income taxation) or gains (under consumption taxation).

According to the traditional approach, the income effects on the taxed bases are not included in the social valuation of one unit of private income. Therefore, the problem with the traditional definition is that it compares the *social* marginal value of public income to the *private* marginal value of private income – up to a transformation of the latter through the social welfare function if non-utilitarian social preferences are adopted. If the marginal cost of public funds is based on

¹⁴The sensitivity of MCF' to the normalization of the tax system was already pointed out by Atkinson and Stern (1974) without using the MCF terminology, however. In particular, MCF' can be shown to be bigger (smaller) than one if consumption goods (factor supplies) are taxed, and factor supplies (consumption goods) are not taxed.

the economically more intuitive definition for the social marginal value of private income based on Diamond (1975), then it correctly compares the *social* marginal value of public income to the *social* marginal value of private income. It will be demonstrated that the traditional definition of the marginal cost of public funds is indeed conceptually problematic, and the three problems identified above are nullified using the Diamond-based approach to the marginal cost of public funds. The social value of private income based on Diamond (1975) is thus defined as follows.

Definition 5 Diamond's definition of the social marginal value of private income – *The social marginal value of one unit of private income accruing to household n equals*

$$\alpha(n) \equiv \Psi'(u(n)) \lambda(n) + \eta t n \frac{\partial l(n)}{\partial T} + \eta \tau \frac{\partial c(n)}{\partial T}, \quad \forall n. \quad (33)$$

The direct utility effect is still measured by $\Psi'(u(n)) \lambda(n)$. In addition, transferring one unit of funds to the household reduces labor supply ($\frac{\partial l(n)}{\partial T} < 0$) and raises consumption demand ($\frac{\partial c(n)}{\partial T} > 0$), since leisure and consumption are assumed to be normal goods. Hence, the government loses $tn \frac{\partial l(n)}{\partial T}$ in tax revenues from the income tax (if $t > 0$) or gains $\tau \frac{\partial c(n)}{\partial T}$ in revenues from the consumption tax (if $\tau > 0$). The welfare effects are obtained by multiplication of the revenue changes with η , the social marginal value of one unit of public resources. Adopting $\alpha(n)$ as the social marginal value of private income makes it possible to define the marginal cost of funds.

Definition 6 Marginal cost of public funds – *The marginal cost of public funds based on the Diamond measure of the social marginal value of income is given by*

$$MCF \equiv \frac{\eta}{\int_{\mathcal{N}} \alpha(n) dF(n)}. \quad (34)$$

Analogously to the traditional marginal cost of public funds, the Diamond-based measure for the marginal cost of funds MCF is defined as the ratio of the social marginal value of public resources η and the average social marginal value of private income $\alpha(n)$. Similarly, the definitions for the distributional characteristics can be adjusted using $\alpha(n)$ instead of $\alpha'(n)$ as the social welfare weights.

Definition 7 Distributional characteristics of the tax bases – *The distributional characteristics ξ_y of tax bases $y(n) = \{z(n), c(n)\}$ based on the Diamond measure of the social marginal value of income are given by*

$$\xi_y \equiv 1 - \frac{\int_{\mathcal{N}} \alpha(n) y(n) dF(n)}{\int_{\mathcal{N}} \alpha(n) dF(n) \int_{\mathcal{N}} y(n) dF(n)} > 0. \quad (35)$$

Definition 8 Distributional characteristic of the public good – *The distributional characteristic of the public good based on the Diamond measure of the social marginal value of income is*

$$\xi_G \equiv 1 - \frac{\int_{\mathcal{N}} \alpha(n) \frac{u_G(\cdot)}{u_c(\cdot)} dF(n)}{\int_{\mathcal{N}} \alpha(n) dF(n) \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n)}. \quad (36)$$

Using the new definitions, the optimal policies under income taxation are expressed as follows.

Proposition 3 *Under the Diamond based MCF definition, and the consumption tax normalized to zero, the optimal rules for public-goods provision and the linear income tax are characterized by*

$$(1 - \xi_G) N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = (1 - \gamma_t \bar{\varepsilon}_{lG}^c) \cdot p, \quad (37)$$

$$MCF_T = 1, \quad (38)$$

$$MCF_t = \frac{1 - \xi_z}{1 + \frac{t}{1-t} \bar{\varepsilon}_{lt}^c} = \frac{1 - \xi_z}{1 - MEB_t}, \quad (39)$$

$$MCF = MCF_T = MCF_t = 1. \quad (40)$$

Proof. Equation (14) can be simplified by setting $\tau = 0$ and substituting equation (34) to find equation (38). Equation (15) is simplified by using equation (35), the Slutsky equation $\frac{\partial l^u(n)}{\partial t} = \frac{\partial l^c(n)}{\partial t} - nl(n) \frac{\partial l(n)}{\partial T}$, and setting $\tau = 0$ to find the first part of equation (39). The second part follows from substituting equation (22). Equation (17) is simplified by using equation (36), the Slutsky equation $\frac{\partial l^u(n)}{\partial G} = \frac{\partial l^c(n)}{\partial G} + \frac{u_G}{\lambda(n)} \frac{\partial l(n)}{\partial T}$, setting $\tau = 0$, and using $\gamma_t \equiv Nt \int_{\mathcal{N}} nl(n) dF(n) / pG$ to find equation (37). ■

Equation (37) reveals that the modified Samuelson rule is almost the same as before, except that the compensated cross-elasticity of labor supply with respect to public goods $\bar{\varepsilon}_{lG}^c$ enters the expression (rather than the uncompensated cross-elasticity). The compensated cross-elasticity is a different measure than the uncompensated cross elasticity. Whereas the uncompensated elasticity is zero with utility functions exhibiting (weak) separability between public goods and labor (leisure), the compensated cross-elasticity is generally different from zero and positive for a wide class of utility functions including the separable ones; see Jacobs (2009). However, the most important difference with proposition 1 is that there should be *no* correction for the marginal cost of funds on the cost side of the modified Samuelson rule. Consequently, tax distortions do not affect the decision rule for the optimal supply of public goods. Although the decision rule does not contain a correction for the marginal cost of funds in second-best, the level of public-goods provision is generally different, since the allocations are not identical in first- and second-best.

Equation (38) demonstrates that, at the optimal tax system, the marginal cost of funds for the lump-sum tax (MCF_T) is always equal to one. The finding that the marginal cost of public funds is equal to one is therefore merely a statement that the tax system is optimal. Intuitively, transferring one unit of resources from the private to the public sector yields exactly one unit in terms of average social benefits. Thus, at the optimum, the government should be indifferent between transferring funds from the public to the private sector. Hence, the first problem that the lump-sum tax does not have a marginal cost of public funds equal to one is resolved.

Equation (39) shows that in Ramsey (1927) analyses without lump-sum taxes and redistributive concerns ($\xi_z = 0$) the marginal cost of public funds for the distortionary tax would be directly related to the marginal excess burden of the tax. Indeed, it is exactly equal to the inverse of $1 - MEB_t$. This finding resolves the second problem that there is no link between the

marginal cost of public funds and the excess burden of taxation. Indeed, by failing to account for the income effects on taxed bases in the social marginal value of private income, these income effects pop up in the expression for the marginal cost of funds; see equation (27).

However, in Mirrlees (1971) analyses, the government only introduces distortionary taxes if this contributes to equality (i.e., if taxing labor income yields distributional benefits). With distributional concerns, $\xi_z > 0$, MCF_t is lowered, as equation (39) reveals. At the optimal tax system, the marginal cost of funds for the income tax should be equal to the marginal cost of funds for the lump-sum tax, as (40) demonstrates. Therefore, the deadweight losses of income taxes should be exactly equal to the distributional gains: $MEB_t = \frac{t}{1-t} \bar{\varepsilon}_{lt}^c = \xi_z$. At the optimum, the marginal excess burden of a distorting tax rate (expressed in monetary units, as a percentage of taxed income) exactly equals the marginal benefits of redistribution (expressed in monetary units, as a percentage of taxed income). Tax distortions are therefore only introduced insofar as these distortions contribute to a more equal distribution of welfare. The more society cares about distribution, the larger is ξ_z and the higher is the optimal income tax. The more elastically labor supply responds to taxes, the larger $-\bar{\varepsilon}_{lt}^c$ will be, and the lower will be the optimal income tax. This is the standard trade-off between equity and efficiency.

Naturally, there should be no correction of the modified Samuelson rule (37) if the public good is financed at the margin with the lump-sum tax, since there is no deadweight loss involved. However, neither should it contain a correction if the marginal source of finance for the public good is the distortionary tax instrument. This is an application of the Envelope Theorem: the deadweight loss of a marginally higher tax rate exactly cancels against the distributional gain if the tax rate is optimally set.

The marginal excess burden of the income tax is therefore the *price of equality* and not the price of public-goods provision. Tax distortions are not introduced for public-goods provision. If the government would not be interested in redistribution, the distributional characteristic would be zero $\xi_z = 0$, and the marginal excess burden would be zero, since distortionary tax rates would be optimally zero. Thus, in the absence of a preference for redistribution, all public goods would be financed with non-distortionary non-individualized taxes.

Correctly specifying the marginal cost of public funds has thus been necessary to prove that the marginal cost of funds should not play a role in the Samuelson rule for optimal linear income taxation. By using the traditional definition one could get the impression that the Samuelson rule contains corrections for both tax distortions and distributional gains associated with taxation, as in Sandmo (1998).

The next proposition demonstrates that all results remain valid also using the different normalization with consumption taxes rather than income taxes.

Proposition 4 *Under the Diamond based MCF definition, and with the income tax normalized to zero, the optimal rules for public-goods provision and the linear consumption tax are*

characterized by

$$(1 - \xi_G) N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = (1 - \gamma_\tau \bar{\varepsilon}_{cG}^c) \cdot p, \quad (41)$$

$$MCF_T = 1, \quad (42)$$

$$MCF_\tau = \frac{1 - \xi_c}{1 + \frac{\tau}{1+\tau} \bar{\varepsilon}_{c\tau}^c} = \frac{1 - \xi_c}{1 - MEB_\tau}, \quad (43)$$

$$MCF = MCF_T = MCF_\tau = 1. \quad (44)$$

Proof. Equation (14) can be simplified by setting $t = 0$ and substituting equation (34) to find equation (42). Equation (15) is simplified by using equation (35), the Slutsky equation $\frac{\partial c^u(n)}{\partial \tau} = \frac{\partial c^c(n)}{\partial \tau} - c(n) \frac{\partial c(n)}{\partial T}$, and setting $t = 0$ to find the first part of equation (43). The second part follows from substituting equation (22). Equation (17) is simplified by using equation (36), the Slutsky equation $\frac{\partial c^u(n)}{\partial G} = \frac{\partial c^c(n)}{\partial G} + \frac{u_G}{\lambda(n)} \frac{\partial c(n)}{\partial T}$, setting $t = 0$, and using $\gamma_\tau \equiv N\tau \int_{\mathcal{N}} c(n) dF(n) / pG$ to find equation (41). ■

Proposition 4 demonstrates that the marginal cost of funds measures have become independent from the particular normalization of the tax system. At the optimal tax system, the marginal cost of funds remains equal to one for all tax instruments, and the excess burden of the consumption tax still equals its distributional gain $\frac{\tau}{1+\tau} \bar{\varepsilon}_{c\tau}^c = \xi_c$. Hence, the characterization of the optimal policy rules has become independent from the particular tax normalization.

The marginal excess burdens of consumption and income taxes – as defined in Lemma 1 – are not *quantitatively* identical under different tax normalizations in the absence of distributional concerns ($\xi_z = \xi_c = 0$), even though allocations are; see also Håkonsen (1998). The contributions by Triest (1990) and Dahlby (2008) are related. If there are no distributional concerns, and lump-sum taxes are ruled out (as in Håkonsen (1998)), then $MCF_t = (1 - MEB_t)^{-1} \neq MCF_\tau = (1 - MEB_\tau)^{-1}$, even if the Diamond concept of the social marginal value of private income is adopted. To resolve this apparent inconsistency, Triest (1990), Håkonsen (1998) and Dahlby (2008) develop alternative MCF concepts relying on correcting the MCF measures for the shadow values of public resources (η), both before and after the introduction of distortionary taxes. However, the explanation for the difference in the MCF measures is that the marginal excess burdens are expressed in terms of different tax bases. At identical allocations, both tax instruments have equal marginal excess burdens in absolute terms, but these excess burdens are expressed as fractions of different tax bases. Hence, the MEB measures differ quantitatively. Once distributional concerns are included, this issue has become moot, since the marginal distributional benefits of each distorting tax instrument are expressed as a fraction of the same tax base as the excess burden of the tax instrument. At the optimal tax system, marginal distortionary costs and marginal distributional gains are equal, and the normalization of the costs or benefits of the tax instrument with a particular tax base becomes immaterial.

The next proposition derives a special case in which the first-best Samuelson rule for public-goods provision is obtained in second-best settings with distortionary taxation.

Proposition 5 *If private utility is given by $u(n) \equiv c(n) - v(l(n)) + \Gamma(G)$, $v', \Gamma' > 0$, $v'' > 0$, $\Gamma'' \leq 0$, then the optimal provision of public goods follows the first-best Samuelson rule in second-*

best settings with optimal distortionary taxation:

$$N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = p. \quad (45)$$

Proof. The first-order condition for labor supply is given by $v'(l(n)) = \frac{(1-t)n}{(1+\tau)}$, $\forall n$. Therefore, $\varepsilon_{lG}^c = 0$. Furthermore, $\frac{u_G}{u_c} = \Gamma'(G)$ is independent from skill n ; hence, $\xi_G = 0$. Substitution of $\varepsilon_{lG}^c = 0$ and $\xi_G = 0$ in (37) or (41) yields the result. ■

The public good does not raise or lower compensated labor supply. Therefore, the public good cannot be used to alleviate the tax distortions on labor supply. Similarly, the public good does not affect the welfare distribution, since every household benefits to an equal extent from the public good. Indeed, the public good is a perfect substitute for the lump-sum cash transfer T . Since the provision of the public good neither affects efficiency nor equity, provision of the public good should follow the first-best policy rule. This simple special case demonstrates why it is misleading to ignore distributional concerns in the analysis of the marginal cost of public funds; the cost side of the Samuelson rule would be multiplied with MCF , whereas the benefits of the distortionary tax would be left out of the picture.

3.3 Sub-optimal taxation

To conclude the section on linear taxation, suppose that the government cannot optimize the lump-sum tax, and has to resort to distortionary income taxation as the marginal source of finance for public goods. The following proposition derives the marginal cost of public funds and the Samuelson rule for public goods provision.

Proposition 6 *Under the Diamond based MCF definition, the lump-sum tax exogenously given, and the consumption tax normalized to zero, the optimal rule for public-goods provision and the marginal cost of public funds are characterized by*

$$(1 - \xi_G) N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = (1 - \gamma_t \bar{\varepsilon}_{lG}^c) \cdot MCF_t \cdot p, \quad (46)$$

$$MCF_t = \frac{1 - \xi_z}{1 + \frac{t}{1-t} \bar{\varepsilon}_{lt}^c} = \frac{1 - \xi_z}{1 - MEB_t} \begin{matrix} \geq \\ \leq \end{matrix} 1. \quad (47)$$

Proof. See Proposition 3. ■

The modified Samuelson rule contains a correction for both the excess burden and the distributional benefits of the distortionary tax that is used to finance the public good. The distributional benefits are omitted in models adopting a representative agent. Clearly, when the lump-sum tax cannot be optimized, the marginal cost of public funds ceases to be equal to one. However, whether it is bigger or smaller than one depends on the distributional and efficiency effects of the tax rate. If the income tax is sub-optimally low (high) from a distributional perspective ($\xi_z > (<) MEB_t$), then the marginal cost of public funds is smaller (bigger) than one, and optimal public goods provision is larger (smaller), everything else equal. If the income tax

is optimized, both the distributional gains and the deadweight losses cancel in the second-best modified Samuelson rule.

4 Optimal non-linear taxation and public-goods provision

The analysis has thus far been confined to linear policy instruments. In the real world, however, most tax systems are non-linear. This section explores the optimal provision of public goods under optimal non-linear taxes; as in Mirrlees (1971).

4.1 Households

Let the non-linear tax schedule be denoted by $T(z(n))$, where $T'(z(n)) \equiv \frac{dT(z(n))}{dz(n)}$ denotes the marginal income tax rate at income $z(n)$. The government observes only the total income of a household ($z(n) = nl(n)$), not its labor supply $l(n)$ or ability n , which rules out individualized lump-sum taxes. The household budget constraint is modified:

$$c(n) = z(n) - T(z(n)), \quad \forall n. \quad (48)$$

The first-order condition for labor supply of each household reads as

$$\frac{-u_l(c(n), l(n), G; n)}{u_c(c(n), l(n), G; n)} = (1 - T'(z(n)))n, \quad \forall n. \quad (49)$$

4.2 Government

The social welfare function remains the same as in (9). In the Mirrlees setup it is more convenient to work with the economy's resource constraint rather than the government budget constraint.¹⁵ The economy's resource constraint is:

$$N \int_{\mathcal{N}} (nl(n) - c(n)) dF(n) = pG. \quad (50)$$

4.3 Incentive compatibility

To determine the non-linear policy schedule $T(\cdot)$, a standard mechanism-design approach is adopted. First, the revelation principle is used to derive the optimal direct mechanism, which induces households to reveal their ability truthfully through self-selection. This direct mechanism yields the optimal second-best allocation. Then, this allocation is decentralized as the outcome of a competitive equilibrium by employing the non-linear policy income tax $T(\cdot)$.

Any second-best allocation must satisfy the incentive-compatibility constraints. Since n is not observable by the government, every bundle $\{c(n), z(n)\}$ for household n must be such that each household self-selects into this bundle and does not prefer another bundle $\{c(m), z(m)\}$

¹⁵If the government maximizes social welfare subject to the resource constraint, and all households respect their budget constraints, then the government budget constraint is automatically satisfied by Walras' law.

intended for household $m \neq n$. By adopting the first-order approach, as in Mirrlees (1971), the incentive constraints can be summarized by a differential equation on utility:

$$\frac{du(c(n), l(n), G; n)}{dn} = u_n(c(n), l(n), G; n) - \frac{l(n)}{n} u_l(c(n), l(n), G; n), \quad \forall n. \quad (51)$$

The first-order approach yields an implementable allocation only if the second-order conditions for utility maximization are fulfilled at the optimum allocation. Lemma 2 states the conditions under which the first-order approach is both necessary and sufficient for a welfare optimum.

Lemma 2 Second-order incentive-compatibility – *The optimal allocation derived under the first-order approach is implementable with the non-linear income tax if i) utility $U(c(n), z(n), G, n) \equiv u\left(c(n), \frac{z(n)}{n}, G; n\right)$ satisfies the Spence-Mirrlees condition; i.e., $\frac{d(-U_z/U_c)}{dn} < 0$, and ii) gross incomes are non-decreasing with skill n ; i.e., $\frac{dz(n)}{dn} \geq 0$.*

Proof. Define the net utility $\Omega(m, n)$ of an n -type mimicking an m -type as $\Omega(m, n) \equiv U(c(m), z(m), G, n) - U(c(n), z(n), G, n)$. Under a direct mechanism, each household reports an ability-type to the government. Each report m is associated with a commodity bundle $\{c(m), z(m)\}$. The optimal utility-maximizing report m by agent n follows from the first-order condition $\frac{d\Omega(m, n)}{dm} = U_c \frac{dc(m)}{dm} + U_z \frac{dz(m)}{dm} = 0$. The second-order condition for this problem is $\frac{\partial^2 \Omega(m, n)}{\partial m^2} \leq 0$. The second-order condition can be simplified. Totally differentiating the first-order condition yields $\frac{\partial^2 \Omega(m, n)}{\partial m^2} + \frac{\partial^2 \Omega(m, n)}{\partial m \partial n} = 0$. Differentiating the first-order condition with respect to n gives $\frac{\partial^2 \Omega(m, n)}{\partial m \partial n} = \frac{\partial U_c}{\partial n} \frac{dc(m)}{dm} + \frac{\partial U_z}{\partial n} \frac{dz(m)}{dm}$. The second-order condition for incentive-compatibility (evaluated at $n = m$) thus requires $\frac{\partial U_c}{\partial n} \frac{dc(n)}{dn} + \frac{\partial U_z}{\partial n} \frac{dz(n)}{dn} \geq 0$. Substitute the first-order condition $\frac{dc(n)}{dn} = -\frac{U_z}{U_c} \frac{dz(n)}{dn}$ to find $\left(\frac{\partial U_z}{\partial n} - \frac{U_z}{U_c} \frac{\partial U_c}{\partial n}\right) \frac{dz(n)}{dn} \geq 0$. Differentiate $\frac{U_z}{U_c}$ to obtain $\frac{d\left(\frac{U_z}{U_c}\right)}{dn} = \frac{1}{U_c} \left(\frac{\partial U_z}{\partial n} - \frac{U_z}{U_c} \frac{\partial U_c}{\partial n}\right)$. Substitute these results to find the final condition: $-\frac{d\left(\frac{-U_z}{U_c}\right)}{dn} \frac{dz(n)}{dn} \geq 0$. Sufficient conditions for second-order incentive-compatibility are therefore: $\frac{d(-U_z/U_c)}{dn} < 0$, and $\frac{dz(n)}{dn} \geq 0$. ■

Lemma 2 implies that the second-order conditions are locally satisfied. It is a standard result in the Mirrlees model without public goods provision, see also ?, Ebert (1992), and ?. Since the public good is identical for all agents, incentive compatibility constraints are not affected by introducing public goods, see also Kreiner and Verdellin (2010). Even if different individuals exhibit a different valuation for the public good, one individual cannot mimic another individual's ability to benefit from the public good, since his utility function remains the same. In the remainder of this paper it's assumed that Lemma 2 holds.

4.4 Optimal taxation and public-goods provision

The Lagrangian for maximizing social welfare is obtained after integrating equation (51) by parts:

$$\begin{aligned} \max_{\{l(n), u(n), G\}} \mathcal{L} &= \int_{\mathcal{N}} (\Psi(u(n)) + \eta [nl(n) - c(n) - pG/N]) f(n) dn \\ &+ \int_{\mathcal{N}} \left(\theta(n) \left(u_l(\cdot) \frac{l(n)}{n} - u_n(\cdot) \right) - u(n) \frac{d\theta(n)}{dn} \right) dn + \theta(\bar{n}) u(\bar{n}) - \theta(\underline{n}) u(\underline{n}). \end{aligned} \quad (52)$$

$\theta(n)$ is the Lagrange multiplier associated with the differential equation for utility in equation (51). η is the Lagrange multiplier on the resource constraint. The transversality conditions for this optimal control problem are given by $\lim_{n \rightarrow \bar{n}} \theta(n) = 0$, and $\lim_{n \rightarrow \underline{n}} \theta(n) = 0$. The first-order conditions for an optimal allocation are – omitting the n -indices –

$$\frac{\partial \mathcal{L}}{\partial l} = \eta \left(n - \frac{dc}{dl} \Big|_{\bar{u}, \bar{G}} \right) f + \frac{\theta u_l}{n} \left(1 + \left(\frac{lu_{ll} - nu_{nl}}{u_l} \right) + \left(\frac{lu_{lc} - nu_{nc}}{u_l} \right) \frac{dc}{dl} \Big|_{\bar{u}, \bar{G}} \right) = 0, \quad \forall n, \quad (53)$$

$$\frac{\partial \mathcal{L}}{\partial u} = \left(\Psi' - \eta \frac{dc}{du} \Big|_{\bar{l}, \bar{G}} \right) f + \theta \left(\frac{lu_{lc}}{n} - u_{nc} \right) \frac{dc}{du} \Big|_{\bar{l}, \bar{G}} - \frac{d\theta}{dn} = 0, \quad \forall n, \quad (54)$$

$$\frac{\partial \mathcal{L}}{\partial G} = \int_{\mathcal{N}} \left(-\eta \left(\frac{p}{N} + \frac{dc}{dG} \Big|_{\bar{l}, \bar{u}} \right) f + \theta \left(\left(\frac{lu_{lc}}{n} - u_{nc} \right) \frac{dc}{dG} \Big|_{\bar{l}, \bar{u}} + \frac{lu_{lG}}{n} - u_{nG} \right) \right) dn = 0. \quad (55)$$

The following proposition provides the conditions for the optimal provision of public goods and the optimal non-linear income tax.

Proposition 7 *The optimal rules for public-goods provision and the non-linear income tax are characterized by*

$$\frac{\theta(n)}{\eta} = \int_n^{\bar{n}} \left(\frac{1}{u_c} - \frac{\Psi'}{\eta} \right) \exp \left[- \int_n^m \left(\frac{\partial \ln u_c}{\partial \ln l} - \frac{\partial \ln u_c}{\partial \ln s} \right) \frac{ds}{s} \right] f(m) dm, \quad \forall n, \quad (56)$$

$$N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = p + N \int_{\mathcal{N}} \frac{\theta(n) u_G(\cdot) / \eta}{n f(n)} \left(\frac{\partial \ln(u_G/u_c)}{\partial \ln n} - \frac{\partial \ln(u_G/u_c)}{\partial \ln l(n)} \right) dF(n), \quad (57)$$

$$\frac{T'(z(n))}{1 - T'(z(n))} = \left(1 + \frac{1}{\varepsilon(n)} \right) \frac{\theta(n) u_c(\cdot) / \eta}{n f(n)}, \quad \forall n, \quad (58)$$

$$\varepsilon(n) \equiv \left(\frac{\partial \ln(-u_l/u_c)}{\partial \ln l} - \frac{\partial \ln(-u_l/u_c)}{\partial \ln n} \right)^{-1} > 0 \quad \forall n \quad (59)$$

Proof. Equation (53) can be rewritten by totally differentiating the utility function at $du = dG = 0$, and substituting the result in the first-order condition for labor supply to find $\frac{dc}{dl} \Big|_{\bar{u}, \bar{G}} = \frac{-u_l}{u_c} = (1 - T')n$. Derive that $\frac{lu_{ll} - nu_{nl}}{u_l} = \frac{\partial \ln u_l}{\partial \ln l} - \frac{\partial \ln u_l}{\partial \ln n}$ and $\frac{lu_{lc} - nu_{nc}}{u_c} = \frac{\partial \ln u_c}{\partial \ln l} - \frac{\partial \ln u_c}{\partial \ln n}$. Substitution of these results and simplifying yields (58). Equation (54) can be rewritten using $\frac{dc}{du} \Big|_{\bar{l}, \bar{G}} = \frac{1}{u_c}$ and $\frac{lu_{lc}}{nu_c} - \frac{u_{nc}}{u_c} = \frac{1}{n} \left(\frac{\partial \ln u_c}{\partial \ln l} - \frac{\partial \ln u_c}{\partial \ln n} \right)$. The resulting equation is a linear differential equation in θ , which can be integrated in a straightforward fashion to find equation (56). Equation (55) is found by totally differentiating utility at $dl = du = 0$ to find $\frac{dc}{dG} \Big|_{\bar{l}, \bar{u}} = -\frac{u_G}{u_c}$. Derive that

$\left(\frac{lu_{lc}}{n} - u_{nc}\right) \frac{dc}{dG} \Big|_{\bar{l}, \bar{u}} + \left(\frac{lu_{lG}}{n} - u_{nG}\right) = \frac{u_G}{n} \left(\frac{\partial \ln(u_G/u_c)}{\partial \ln l} - \frac{\partial \ln(u_G/u_c)}{\partial \ln n} \right)$, and rearrange to find equation (57). ■

Equation (56) yields the solution of $\theta(n)/\eta$, which equals the marginal increase in social welfare (in monetary units) of redistribution of one unit of income from households above n to households below n . $\theta(n) = 0$ everywhere if the government is not interested in redistribution.

Equation (57) gives the policy rule for the optimal provision of public goods. Define

$$\Delta(n) \equiv \frac{T'(z(n))}{1 - T'(z(n))} \left(1 + \frac{1}{\varepsilon(n)}\right)^{-1} \left(\frac{\partial \ln(u_G/u_c)}{\partial \ln l(n)} - \frac{\partial \ln(u_G/u_c)}{\partial \ln n} \right), \quad \forall n, \quad (60)$$

and use the first-order condition for the income tax (58), to rewrite (57) as

$$N \int_{\mathcal{N}} (1 + \Delta(n)) \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = p. \quad (61)$$

$\Delta(n)$ indicates the extent to which public goods are overprovided relative to the first-best Samuelson rule; cf. equation (45). Whether public goods are under- or overprovided compared to the first-best rule is determined by the presence of the distortionary income tax ($T'(z(n)) > 0$). The more distortionary is income taxation, as indicated by a larger elasticity $\varepsilon(n)$, the more public-goods provision deviates from the first-best policy rule. $\Delta(n)$ can be interpreted as an implicit subsidy on public-goods provision at skill level n . $\Delta(n)$ is different from zero if the marginal willingness to pay for the public good $\frac{u_G(\cdot)}{u_c(\cdot)}$ varies with labor supply or with ability. In the absence of redistributive concerns, the marginal value of redistribution is nil ($\theta(n) = 0$), marginal tax rates are zero ($T'(z(n)) = 0$), and so is the implicit subsidy on public-goods provision ($\Delta(n) = 0$). In that case, public-goods provision satisfies the first-best Samuelson rule (45).

The public good is more (less) complementary to work than consumption goods are, if the marginal willingness to pay for the public good rises (falls) with labor effort (i.e., $\frac{\partial \ln(u_G/u_c)}{\partial \ln l} > 0 (< 0)$).¹⁶ The provision of public goods then increases. The intuition is the same as in Atkinson and Stiglitz (1976). The government should provide more public goods if they are more complementary to work than private goods are, and provide fewer public goods if they are more complementary to leisure than private goods are. In doing so, the government alleviates the distortions of the labor income tax on work effort. To put it differently, if $\frac{\partial \ln(u_G/u_c)}{\partial \ln l} > 0 (< 0)$, a higher (lower) level of public-goods provision relaxes the incentive constraints associated with redistribution of income, as households with high ability are less tempted to mimic households with a lower ability. The Atkinson and Stern (1974) term capturing the interaction of the public goods with labor supply is therefore also present under non-linear taxation. See also Christiansen (1981) and Boadway and Keen (1993).

¹⁶The term $\frac{\partial \ln(u_G/u_c)}{\partial \ln l(n)}$ can be rewritten as $\omega_l (\rho_{lG} - \rho_{lc})$, where $\omega_l \equiv \frac{-lu_l}{u} > 0$ is the utility share of labor, $\rho_{lG} \equiv \frac{u_{lG}u}{(-u_l)u_G}$ is the Hicksian partial elasticity of complementarity between public goods and labor, and $\rho_{lc} \equiv \frac{u_{lc}u}{(-u_l)u_c}$ is the Hicksian partial elasticity of complementarity between private goods and labor. Thus, the willingness to pay for the public good increases with labor supply if $\rho_{lG} > \rho_{lc}$. That is, if public goods are stronger Hicksian complements to labor than private goods are.

When the utility function differs across households, public-goods provision should also be employed for redistribution. Public-goods provision can extract additional information on the skill level, since the willingness to pay for the public good directly varies with skills. If the marginal willingness to pay for the public good rises (falls) with ability (i.e., $\frac{\partial \ln(u_G/u_c)}{\partial \ln n} > 0$ (< 0)), then the public good benefits the households with higher skill levels relatively more (less).¹⁷ Consequently, the level of public-goods provision falls (increases) for redistributive reasons. This is analogous to the presence of the ξ_G -term in the modified Samuelson rule obtained under the linear case. If $\frac{\partial \ln(u_G/u_c)}{\partial \ln n} > 0$ (< 0), then high-skill types are less (more) tempted to mimic low-skill types if public-goods provision expands, since they have a stronger (weaker) preference for public goods. Consequently, incentive-compatibility constraints are relaxed (tightened), and public-goods provision reduces (increases) the distortions of redistributing income. If the utility function is identical for all households n , this term drops out (i.e., $\frac{\partial \ln(u_G/u_c)}{\partial \ln n} = 0$). Intuitively, differences in earning ability are then the only source of inequality. If the marginal willingness to pay for the public good does not directly vary with ability, the government is unable to extract more information on hidden ability by providing public goods – *ceteris paribus*. Therefore, the public good is not under- or overprovided relative to the first-best rule for redistributive reasons. Indeed, the non-linear income tax is then a more efficient instrument for redistributing income than is public-goods provision. See also Christiansen (1981) and Boadway and Keen (1993).

This finding contrasts with the case of the linear income tax; optimal provision of the public good is *always* found to be dependent on the distributional impact of the public good, except for the (trivial) case in which the marginal willingness to pay is equal for all households. With linear taxation, the government uses an informationally inferior instrument to redistribute income by ignoring the information on individual earnings. Hence, the government optimally complements the income tax by using indirect instruments for redistribution, such as public-goods provision.

The expression for the optimal non-linear income tax in equation (58) is identical to Mirrlees (1971). $\varepsilon(n) > 0$ is a measure for the compensated labor supply elasticity, see also equation (59). At each point in the income distribution, the government trades off the distortions of a higher marginal tax rate against the distributional gains of the marginal tax. If ε is larger, the optimal tax rate should be set lower. When households have identical preferences, $\varepsilon \equiv \frac{\partial \ln l}{\partial \ln((1-T)n)}$ is obtained, which exactly equals the compensated labor supply elasticity. If households with a higher ability have a lower willingness to pay for leisure, then $\frac{\partial \ln(-u_l/u_c)}{\partial \ln n} < 0$, and the compensated labor supply elasticity falls as ability increases (and vice versa).

If the taxed labor base is larger, then $nf(n)$ is higher, and the welfare losses of distortionary taxes increase, because more households or households with higher skill levels are affected by higher marginal tax rates. The marginal tax rates should be higher if the marginal value of redistribution $\frac{\theta(n)}{\eta}$ increases. The marginal value of redistribution is zero at the bottom of the

¹⁷The term $\frac{\partial \ln(u_G/u_c)}{\partial \ln n}$ can be rewritten as $\omega_n (\rho_{nG} - \rho_{nc})$, where $\omega_n \equiv \frac{nu_n}{u} > 0$ is the utility share of ability, $\rho_{nG} \equiv \frac{u_{nG}u}{u_n u_G}$ is the Hicksian partial elasticity of complementarity between public goods and ability, and $\rho_{nc} \equiv \frac{u_{nc}u}{u_n u_c}$ is the Hicksian partial elasticity of complementarity between private goods and ability. The willingness to pay for the public good increases with ability if $\rho_{nG} > \rho_{nc}$. That is, if public goods are stronger Hicksian complements to ability than private goods are.

income distribution (in the absence of bunching at zero labor effort). Hence, marginal tax rates are zero. The intuition is that it is not optimal to distort labor supply for the lowest-skilled household if there is no one below that household benefitting from redistribution. If there is bunching at zero labor supply, the marginal tax rate at the bottom is positive, since positive marginal tax rates then help to redistribute income from households with positive labor supply to households with zero income; see Seade (1977) and Ebert (1992). Similarly, if the skill distribution is bounded, the marginal tax rate should be zero at the top. Intuitively, it makes no sense to distort the labor supply of the highest-skilled household, since there will be no one contributing to more redistribution (cf. Sadka (1976) and Seade (1977)). However, if the skill distribution is unbounded, the top rate might not converge to zero; see also Diamond (1998) and Saez (2001). The reader is referred to Mirrlees (1971), Seade (1977), Tuomala (1984), Diamond (1998) and Saez (2001) for more elaborate discussions of the optimal non-linear tax.

Christiansen (1981), Boadway and Keen (1993) and Kaplow (1996) discuss the case in which the utility function is identical for all households and is (weakly) separable between (private/public) consumption and leisure. They find that the optimal provision of public goods then follows the first-best Samuelson rule. Naturally, the allocations will differ between first- and second-best.

Proposition 8 *If private utility is given by $u(n) \equiv u(h(c(n), G), l(n))$, $\forall n$, then the optimal provision of public goods follows the first-best Samuelson rule in second-best settings with optimal distortionary taxation:*

$$N \int_{\mathcal{N}} \frac{u_G(\cdot)}{u_c(\cdot)} dF(n) = p. \quad (62)$$

Proof. If $u \equiv u(h(c, G), l)$, it is immediately established that $\frac{\partial \ln(u_G/u_c)}{\partial \ln l} = \frac{\partial \ln(u_G/u_c)}{\partial \ln n} = 0$. Substitution in (57) yields (62). ■

The reason for this result is that public goods do not help to reduce the tax wedge on labor supply, because the marginal willingness to pay for the public good does not vary with labor supply. Therefore, public-goods provision does not affect the incentive-compatibility constraints. Moreover, it does not help to improve income redistribution, since the marginal willingness to pay for the public good does not vary with skills. As a result, it is not optimal to provide public goods above or below levels that would be prescribed by the first-best Samuelson rule.

4.5 Marginal cost of public funds

The non-linear income tax does not deliver a direct measure for the marginal cost of funds. To determine the marginal cost of funds, the welfare effect is determined of a unit increase in the intercept of the tax function, $-T(0)$. Each tax payer is then provided with one unit of extra income. This is equivalent to an increase in the lump-sum transfer by one unit of income under the linear tax schedule. To determine the welfare effect of this policy, a perturbation of $-T(0)$ along the optimal non-linear tax schedule is analyzed. This method is also employed by Saez (2001) and Jacquet, Lehmann, and Van der Linden (2010).

Proposition 9 *The marginal cost of public funds is equal to one under optimal non-linear income taxation:*

$$MCF \equiv \frac{\eta}{\int_{\mathcal{N}} \left(\Psi'(u(n)) \lambda(n) + \eta n T'(z(n)) \frac{\partial l(n)}{\partial (-T(0))} \right) dF(n)} = 1. \quad (63)$$

Proof. Raising $-T(0)$ by one unit has the following three effects. Firstly, the government loses one unit in resources per capita. Secondly, this policy mechanically raises social welfare – measured in monetary equivalents – for each household by $\frac{\Psi'(u(n))\lambda(n)}{\eta}$. Thirdly, a larger $-T(0)$ generates behavioral changes on labor supply. Since marginal tax rates are unaffected by this policy reform, substitution effects are zero and only the income effects matter. The income effect on labor supply changes tax revenues by $T'(z(n)) \frac{\partial z(n)}{\partial (-T(0))} = T'(z(n)) n \frac{\partial l(n)}{\partial (-T(0))}$, $\forall n$. The total change in social welfare should be zero when the intercept of the tax function is optimized:

$$N \int_{\mathcal{N}} \left(\frac{\Psi'(u(n)) \lambda(n)}{\eta} + n T'(z(n)) \frac{\partial l(n)}{\partial (-T(0))} \right) dF(n) = N. \quad (64)$$

Rewriting yields the desired result. ■

At the optimum, the marginal cost of public funds for all tax instruments must be equalized. Since $-T(0)$ is equivalent to a non-distortionary non-individualized lump-sum tax, at the optimum the marginal cost of funds for all other tax rates must be equal to the marginal cost of funds of $-T(0)$. Thus, at each point in the income distribution, tax distortions should be equal to distributional gains for all marginal tax rates.

To summarize, the analysis of optimal non-linear taxation and public-goods provision qualitatively confirms the findings under optimal linear taxation. Tax distortions are *not* an ingredient of the modified Samuelson rule for the optimal provision of public goods. However, the modified Samuelson rule indicates a role for second-best interactions of public goods with labor supply and/or distributional effects of public goods if the willingness to pay for public goods varies with labor supply and/or ability.

4.6 Normalization tax system

First-order conditions (53), (54) and (55), jointly with the household budget constraints (48), and the economy's resource constraint (50), describe the complete second-best optimal allocation in the economy. However, none of these conditions reveal anything about the particular mechanism by which the optimal allocation should be implemented. It is easily established that a non-linear consumption tax $\tau(c(n))$, where $\tau'(c(n)) \equiv \frac{d\tau(c(n))}{dc(n)}$ denotes the marginal consumption tax rate at consumption level $c(n)$, can also implement the same optimal allocation as the non-linear income tax if optimal consumption tax satisfies $\tau'(c(n)) = \left(1 + \frac{1}{\varepsilon(n)}\right) \frac{\theta(n)\lambda(n)/\eta}{nf(n)}$. The equations for $\theta(n)/\eta$ (56) and the modified Samuelson rule (57), remain the same. Therefore, results for public-goods provision are insensitive to the normalization of the tax system.¹⁸

¹⁸Boadway and Keen (1993) show that there is a numéraire issue under non-linear taxation. Based on Tuomala (1990), they show that when the government uses a non-linear commodity tax on consumption, rather than on

5 Discussion

Saying that the marginal cost of funds equals one is merely making a statement about the optimality of the tax system; if the marginal cost of public funds is one, welfare cannot be increased by reallocating resources between the public and private sector. This result should probably hold more generally in any model analyzing optimal policies allowing for non-distortionary transfers between the public and private sector. This section explores some potential criticisms that could be raised on the result that the marginal cost of public funds is one at the optimal tax system.

5.1 Optimality of the tax system

This paper followed common practice in the second-best literature by assuming that the government is a benevolent social planner and that markets are Pareto-efficient in the absence of government intervention. One can ask whether the government sets taxes optimally and whether markets are always Pareto-efficient in the absence of government intervention.

Benevolent planners may fail to implement second-best optimal policies in the presence of commitment problems; see, for example, Kydland and Prescott (1977) and Fischer (1980). Moreover, in the real world tax systems are not set by benevolent planners, but by politicians. Political systems create various distortions; see, for example, Persson and Tabellini (2000). In addition, there can be all kinds of legal constraints that rule out certain policy instruments. Government failure could explain why taxes and public goods are not set at second-best optimal levels. One may therefore be tempted to conclude that government failure raises the marginal cost of public funds above one. However, this conclusion cannot be drawn in general if the underlying sources of government failure are not explicitly specified. A well-functioning democracy, a just legal system, the presence of countervailing powers, and the protection of citizens against abuses of government power have social benefits. Economic analysis should take into account both the costs and benefits of political and legal systems. It could then be conjectured that the marginal cost of public funds is one if political and legal institutions are optimized; the marginal costs of a given political or legal rule should be equal to its benefits.

Not only governments, but also markets can fail. Asymmetric information, transaction costs, monopoly/monopsony power, and externalities prevent an efficient market outcome. Again, one might be tempted to conclude that market failure pushes the marginal cost of public funds below one, so that government intervention becomes more attractive. However, without explicitly incorporating the fundamental reasons why markets fail into the analysis, it is unwarranted to conclude that the marginal cost of funds lies below one if markets fail. Governments could be confronted with the same fundamental problems that give rise to market failures. Furthermore, from the theory of the second-best, it is well-known that the implications of market failure for the determination of second-best policies are generally ambiguous as government intervention could either exacerbate or reduce pre-existing market failures. For future research, the marginal

income, the preferences under which the standard first-best Samuelson rule applies are different than the weakly separable structure of preferences as in $u_n = u(h(c_n, G), l_n)$. This does not imply, however, that the modified Samuelson rule under non-linear taxation needs to be modified for the marginal cost of funds.

cost of funds should be derived by explicitly modeling market or government failure from first principles, not by introducing ad hoc constraints on the government instrument set.

5.2 Instrument set

Any measure for the marginal cost of public funds obtained from models with ad hoc constraints on the instrument set of the government is bound to be misleading. This paper has demonstrated that ruling out non-individualized lump-sum taxes – as a short-cut for an unspecified distributional problem – has large consequences for the marginal cost of funds. However, one can debate whether, at the margin, governments have access to non-distortionary transfers, given practical (or even legal) problems in implementing lump-sum taxes.

From a theoretical point of view, the lump-sum part of the tax system analyzed in the current paper generally consists of subsidies (i.e., transfers) to households. In particular, the intercept of the tax function is negative. Therefore, it is perfectly feasible to marginally reduce lump-sum transfers to the private sector. Indeed, within the context of the model, there is nothing preventing the government to implement a tax system with lower transfers. Moreover, also in the real world, tax systems generally entail a general tax credit or tax exemption, which can be (and often is) changed by policy makers, without running into practical or legal problems.

One could extend the model to allow for migration of households, resulting in participation constraints as in, for example, ?. This is a large deviation from the assumptions used in the standard literature on the marginal cost of public funds. Nevertheless, even lump-sum taxes/transfers are probably no longer neutral from an efficiency point of view. In particular, a marginal reduction in the transfer induces some individuals to migrate. It could be conjectured that the marginal cost of public funds is then no longer one at the optimal tax system. Future research could explore the role of participation constraints for the marginal cost of public funds.

5.3 Kaplow and sub-optimal taxation

Kaplow (1996), Kaplow (2004) and Kaplow (2008) argue that even when the tax system is not optimal, neither distributional concerns nor incentive effects of the financing of public goods with distortionary taxes should be included in the discussion of second-best policy analysis. Laroque (2005) and Gauthier and Laroque (2009) contain similar ideas and provide rigorous proofs of this claim. The message of Kaplow is, loosely speaking, equivalent to saying that the marginal cost of funds is always equal to one – even in suboptimal tax systems. However, in this argument, the government simultaneously does two things: i) it changes the provision of the public good, and ii) it applies a benefit-absorbing change in the tax schedule that fully extracts households' willingness to pay for the public good. Given that preferences are homogeneous and separable, this tax change does not affect labor supply incentives. Hence, the tax adjustment can perfectly imitate a pure benefit tax. Consequently, the traditional Samuelson rule can be used to judge whether public-goods provision should increase or not. This paper's results are more general in the sense that they apply under non-separable and heterogeneous preferences and also apply to optimal linear taxation. Moreover, a benefit-absorbing tax change is not required, either,

in order to demonstrate that the marginal cost of public funds should not be present in the cost-benefit rule for public-goods provision. However, the important requirement is that the tax system is optimized. A marginal change in distortionary taxes then produces exactly offsetting distortions and distributional gains from taxation that cancel out in the modified Samuelson rule.

5.4 Implications for other public policies

The results of this paper have larger relevance to other public policies than public-goods provision. Sandmo (1975), Bovenberg and de Mooij (1994) and ? show that the tax distortions result in a less aggressive corrective tax policy. Intuitively, corrective taxes exacerbate preexisting distortions of the labor tax. However, these papers do not motivate tax distortions from redistributional concerns. By explicitly incorporating optimal redistribution, Jacobs and de Mooij (2010) demonstrate that the marginal cost of public funds is again equal to one. Thus, it is generally incorrect to set optimal corrective taxes below Pigouvian levels, as in Sandmo (1975), Bovenberg and de Mooij (1994) and ?.

Barro (1979) and Lucas and Stokey (1983) show that under distortionary taxation Ricardian equivalence breaks down, and debt-financing and tax-financing of public spending cease to be equivalent. Since tax distortions are convex in the tax rates, it is better to smooth taxes over time, rather than having time-varying tax rates. Therefore, the appropriate use of public debt sustains tax smoothing. Werning (2007), however, demonstrates that ignoring distributional concerns to motivate tax distortions has fundamental consequences for the theory of optimal debt management. The reason is that, on the margin, the government always has access to non-distortionary sources of finance (i.e., non-individualized lump-sum transfers). Thus, the marginal cost of public funds is constant and equal to one over time. Debt and (lump-sum) tax financing are therefore equivalent and Ricardian equivalence is restored.

Finally, when the marginal cost of funds is equal to one, revenue-raising instruments (such as taxes) are not superior to revenue-neutral instruments (such as regulation), or to revenue-reducing instruments (such as subsidies). The reason is that the fundamental informational constraint – ability is private information – cannot be relaxed by simply adopting a different policy instrument. Quantity controls, regulation or auctions can sustain a second-best optimal allocation equally as well as price instruments. Hence, one needs to include additional constraints into the analysis in order to assess the desirability of revenue-raising over revenue-neutral or revenue-reducing policy instruments.

5.5 Applied cost-benefit analysis

The analysis in this paper shows that one should not make corrections for tax-distortions in the applied cost-benefit analysis if one assumes that taxes are optimally set. Alternatively, if one assumes that taxes are not optimally set, one should include not only the costs of taxation, but also the (distributional) benefits of taxation in the cost-benefit analysis. The valuation of the distributional benefits of taxation is, however, a hazardous task for any policy analyst, since it

inevitably involves a judgement about the desirability of redistribution of welfare. For practical policy analysis it is probably best to invoke Becker (1983)'s efficient redistribution hypothesis, which argues that the political system should achieve an outcome in which all opportunities for political gains to redistribute incomes are exhausted. In that case, the deadweight losses should again be equal to the distributional gains of distortionary taxes, and no correction for the deadweight losses of taxation should be applied in cost-benefit studies.

6 Conclusions

This paper analyzed the simultaneous setting of optimal taxes and the provision of public goods in models with heterogeneous agents having heterogeneous preferences. Both optimal linear and non-linear tax schedules are analyzed. It has been demonstrated that tax distortions are the ultimate result of redistributive concerns. Using the appropriate definition for the social value of private income, the marginal cost of funds at the optimal tax system is one under optimal linear and non-linear taxes. Indeed, saying that the marginal cost of public funds equals one is a making a mere statement of the optimality of the tax system; society cannot gain welfare by reallocation of resources between the public and the private sector, since at the optimum these resources are valued equally by the government and the private sector. Moreover, at the optimal tax system, the marginal cost of public funds for all tax instruments should be equalized. Hence, the marginal cost of funds for distortionary taxation equals the marginal cost of funds for non-distortionary taxation. The distributional benefits of distortionary tax rates should therefore be equal to the marginal excess burden of tax rates.

The modified Samuelson rule should thus *not* be corrected for the marginal cost of funds under linear and non-linear taxation. If the government was not interested in redistribution, then all public goods would be financed with non-distortionary lump-sum taxes. However, there might well be corrections in the Samuelson rule for interactions of the public good with labor supply and the distributional impact of the public good. With specific assumptions on preferences, the first-best Samuelson rule is applicable in the second-best, although second-best allocations naturally differ from first-best allocations. In particular, with optimal linear taxes, utility should be homogeneous across agents, quasi-linear in consumption, and completely separable in its arguments. With non-linear income taxes, utility should be homogeneous and feature weak separability between labor, on the one hand, and consumption of private and public goods, on the other.

References

- Allgood, Sam and Arthur Snow. 1998. "The Marginal Cost of Raising Tax Revenue and Redistributing Income." *Journal of Political Economy* 106 (6):1264–1273.
- Atkinson, Anthony B. and Nicholas H. Stern. 1974. "Pigou, Taxation and Public Goods." *Review of Economic Studies* 41 (1):119–128.

- Atkinson, Anthony B. and Joseph E. Stiglitz. 1976. "The Design of Tax Structure: Direct versus Indirect Taxation." *Journal of Public Economics* 6 (1-2):55–75.
- Ballard, Charles L. and Don Fullerton. 1992. "Distortionary Taxes and the Provision of Public Goods." *Journal of Economic Perspectives* 6 (3):117–131.
- Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy* 87 (5):940–971.
- Becker, Gary S. 1983. "A Theory of Competition Among Pressure Groups for Political Influence." *Quarterly Journal of Economics* 98 (3):371–400.
- Blundell, Richard and Thomas MaCurdy. 1999. "Labor Supply: A Review of Alternative Approaches." In *Handbook of Labor Economics*, vol. 3, edited by Orley Ashenfelter and David Card, chap. 27. Elsevier, 1559–1695.
- Boadway, Robin and Michael Keen. 1993. "Public Goods, Self-Selection and Optimal Income Taxation." *International Economic Review* 34 (3):463–478.
- Bovenberg, A. Lans and Ruud A. de Mooij. 1994. "Environmental Levies and Distortionary Taxation." *American Economic Review* 84 (4):1085–1085.
- Browning, Edgar K. 1976. "The Marginal Cost of Public Funds." *Journal of Political Economy* 84 (2):283–298.
- . 1987. "On the Marginal Welfare Cost of Taxation." *American Economic Review* 77 (1):11–23.
- Browning, Edgar K. and William R. Johnson. 1984. "The Trade-Off Between Equity and Efficiency." *Journal of Political Economy* 92 (2):175–203.
- Christiansen, Vidar. 1981. "Evaluation of Public Projects under Optimal Taxation." *Review of Economic Studies* 48 (3):447–57.
- Dahlby, Bev. 2008. *The Marginal Cost of Public Funds: Theory and Applications*. Cambridge-MA: MIT Press.
- Diamond, Peter A. 1975. "A Many-Person Ramsey Tax Rule." *Journal of Public Economics* 4 (4):335–342.
- . 1998. "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates." *American Economic Review* 88 (1):83–95.
- Diamond, Peter A. and James A. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency." *American Economic Review* 61 (1):8–27.
- Ebert, Udo. 1992. "A Reexamination of the Optimal Nonlinear Income Tax." *Journal of Public Economics* 49 (1):47–73.

- Feldstein, Martin. 1972. "Distributional Equity and the Optimal Structure of Public Prices." *American Economic Review* 62 (1):32–36.
- Fischer, Stanley. 1980. "Dynamic Inconsistency, Cooperation, and the Benevolent Disassembling Government." *Journal of Economic Dynamics and Control* 2:93–107.
- Gauthier, Stephane and Guy R. Laroque. 2009. "Separability and Public Finance." *Journal of Public Economics* 93 (11-12):1168–1174.
- Håkonsen, Lars. 1998. "An Investigation into Alternative Representations of the Marginal Cost of Public Funds." *International Tax and Public Finance* 5 (3):329–343.
- Harberger, Arnold C. 1964. "Taxation, Resource Allocation, and Welfare." In *The Role of Direct and Indirect Taxes in the Federal Revenue System*. Princeton University Press for NBER and the Brookings Institution, 25–70.
- Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 94 (1-2):114–128.
- Jacobs, Bas. 2009. "The Marginal Cost of Public Funds and Optimal Second-Best Policy Rules." Erasmus University Rotterdam.
- Jacobs, Bas and Ruud A. de Mooij. 2010. "Pigou Meets Mirrlees: On the Irrelevance of Tax Distortions for the Optimal Corrective Tax." Erasmus University Rotterdam.
- Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden. 2010. "Optimal Redistributive Taxation with both Extensive and Intensive Responses." IZA Discussion Papers No. 4837.
- Kaplow, Louis. 1996. "The Optimal Supply of Public Goods and the Distortionary Cost of Taxation." *National Tax Journal* 49 (4):513–533.
- . 2004. "On the (Ir)relevance of Distribution and Labor Supply Distortion to Public Goods Provision and Regulation." *Journal of Economic Perspectives* 18 (4):159–175.
- . 2008. *The Theory of Taxation and Public Economics*. Princeton and Oxford: Princeton University Press.
- Kreiner, Klaus Thustrup and Nicolai Verdellin. 2010. "Optimal Provision of Public Goods: A Synthesis." *Scandinavian Journal of Economics* Forthcoming.
- Kydland, Finn E. and Edward C. Prescott. 1977. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy* 85 (3):473–492.
- Laffont, Jean-Jacques and Jean Tirole. 1993. *A Theory of Incentives in Procurement and Regulation*. Cambridge-MA: MIT-Press.

- Laroque, Guy R. 2005. "Indirect Taxation is Superfluous under Separability and Taste Homogeneity: A Simple Proof." *Economics Letters* 87 (1):141–144.
- Lucas, Robert E. jr. and Nancy Stokey. 1983. "Optimal Fiscal and Monetary Policy in an Economy without Capital." *Journal of Monetary Economics* 12 (1):55–93.
- Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Review of Economic Studies* 38 (2):175–208.
- Persson, Torsten and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge-MA: MIT-Press.
- Pigou, Arthur C. 1947. *A Study in Public Finance - 3rd Edition*. London: Macmillan.
- Pirttilä, Jukka and Matti Tuomala. 2001. "On Optimal Non-Linear Taxation and Public Good Provision in an Overlapping Generations Economy." *Journal of Public Economics* 79 (3):485–501.
- Ramsey, Frank P. 1927. "A Contribution to the Theory of Taxation." *Economic Journal* 37 (145):47–61.
- Sadka, Efraim. 1976. "On Income Distribution, Incentive Effects and Optimal Income Taxation." *Review of Economic Studies* 43 (2):261–267.
- Saez, Emmanuel. 2001. "Using Elasticities to Derive Optimal Income Tax Rates." *Review of Economic Studies* 68 (1):205–229.
- . 2002. "The Desirability of Commodity Taxation Under Non-Linear Income Taxation and Heterogeneous Tastes." *Journal of Public Economics* 83 (2):217–230.
- Samuelson, Paul A. 1954. "The Pure Theory of Public Expenditure." *Review of Economics and Statistics* 36 (4):387–389.
- Sandmo, Agnar. 1975. "Optimal Taxation in the Presence of Externalities." *Scandinavian Journal of Economics* 77 (1):86–98.
- . 1998. "Redistribution and the Marginal Cost of Public Funds." *Journal of Public Economics* 70:365–382.
- Seade, Jesus K. 1977. "On the Shape of Optimal Tax Schedules." *Journal of Public Economics* 7 (2):203–235.
- Sheshinski, Eytan. 1972. "The Optimal Linear Income Tax." *Review of Economic Studies* 39 (3):297–302.
- Slemrod, Joel B. and Schlomo Yitzhaki. 2001. "Integrating Expenditure and Tax Decisions: The Marginal Cost of Funds and the Marginal Benefit of Projects." *National Tax Journal* 52 (2):189–201.

- Stiglitz, Joseph E. and Partha Dasgupta. 1971. "Differential Taxation, Public Goods and Economic Efficiency." *Review of Economic Studies* 38 (2):151–174.
- Triest, Robert K. 1990. "The Relationship between the Marginal Cost of Public Funds and Marginal Excess Burden." *American Economic Review* 80 (3):557–566.
- Tuomala, Matti. 1984. "On the Optimal Income Taxation: Some Further Numerical Results." *Journal of Public Economics* 23 (3):351–366.
- . 1990. *Optimal Income Tax and Redistribution*. Oxford: Clarendon Press.
- Werning, Ivan. 2007. "Optimal Fiscal Policy with Redistribution." *Quarterly Journal of Economics* 122 (3):925–967.
- Wildasin, David E. 1984. "On Public Good Provision with Distortionary Taxation." *Economic Inquiry* 22 (2):227–243.
- Wilson, John D. 1991. "Optimal Public Good Provision with Limited Lump-sum Taxation." *American Economic Review* 81 (1):153–166.

WORKING PAPERS

Uppsala Center for Fiscal Studies

Editor: Håkan Selin

- 2009:1 Sören Blomquist and Håkan Selin, Hourly Wage Rate and Taxable Labor Income Responsiveness to Changes in Marginal Tax Rates. 31 pp.
- 2009:2 Luca Micheletto, Optimal nonlinear redistributive taxation and public good provision in an economy with Veblen effects. 26 pp.
- 2009:3 Håkan Selin, The Rise in Female Employment and the Role of Tax Incentives. An Empirical Analysis of the Swedish Individual Tax Reform of 1971. 38 pp.
- 2009:4 Håkan Selin, Marginal tax rates and tax-favoured pension savings of the self-employed Evidence from Sweden. 32 pp.
- 2009:5 Tobias Lindhe and Jan Södersten, Dividend taxation, share repurchases and the equity trap. 27 pp.
- 2009:6 Che-Yan Liang, Nonparametric Structural Estimation of Labor Supply in the Presence of Censoring. 48 pp.
- 2009:7 Sören Blomquist, Vidar Christiansen and Luca Micheletto, Public Provision of Private Goods and Nondistortionary Marginal Tax Rates: Some further Results. 42 pp.
- 2009:8 Laurent Simula and Alain Trannoy, Optimal Income Tax under the Threat of Migration by Top-Income Earners. 26 pp.
- 2009:9 Laurent Simula and Alain Trannoy, Shall We Keep Highly Skilled at Home? The Optimal Income Tax Perspective. 26 pp.
- 2009:10 Michael Neugart and Henry Ohlsson, Economic incentives and the timing of births: Evidence from the German parental benefit reform 2007, 21 pp.
- 2009:11 Laurent Simula, Optimal Nonlinear Income Tax and Nonlinear Pricing: Optimality Conditions and Comparative Static Properties, 25 pp.
- 2009:12 Ali Sina Onder and Herwig Schlunk, Elderly Migration, State Taxes, and What They Reveal, 26 pp.
- 2009:13 Ohlsson, Henry, The legacy of the Swedish gift and inheritance tax, 1884-2004, 26 pp.
- 2009:14 Onder, Ali Sina, Capital Tax Competition When Monetary Competition is Present, 29 pp.

- 2010:1 Sören Blomquist and Laurent Simula, Marginal Deadweight Loss when the Income Tax is Nonlinear. 21 pp.
- 2010:2 Marcus Eliason and Henry Ohlsson, Timing of death and the repeal of the Swedish inheritance tax. 29 pp.
- 2010:3 Mikael Elinder, Oscar Erixson and Henry Ohlsson, The Effect of Inheritance Receipt on Labor and Capital Income: Evidence from Swedish Panel Data. 28 pp.
- 2010:4 Jan Södersten and Tobias Lindhe, The Norwegian Shareholder Tax Reconsidered. 21 pp.
- 2010:5 Anna Persson and Ulrika Vikman, Dynamic effects of mandatory activation of welfare participants. 37 pp.
- 2010:6 Ulrika Vikman, Does Providing Childcare to Unemployed Affect Unemployment Duration? 43 pp.
- 2010:7 Per Engström, Bling Bling Taxation and the Fiscal Virtues of Hip Hop. 12 pp.
- 2010:8 Niclas Berggren and Mikael Elinder, Is tolerance good or bad for growth? 34 pp.
- 2010:9 Heléne Lundqvist, Matz Dahlberg and Eva Mörk, Stimulating Local Public Employment: Do General Grants Work? 35pp.
- 2010:10 Marcus Jacob and Martin Jacob, Taxation, Dividends, and Share Repurchases: Taking Evidence Global. 42 pp.
- 2010:11 Thomas Aronsson and Sören Blomquist, The Standard Deviation of Life-Length, Retirement Incentives, and Optimal Pension Design. 24 pp.
- 2010:12 Spencer Bastani, Sören Blomquist and Luca Micheletto, The Welfare Gains of Age Related Optimal Income Taxation. 48 pp.
- 2010:13 Adrian Adermon and Che-Yuan Liang, Piracy, Music, and Movies: A Natural Experiment. 23 pp.
- 2010:14 Spencer Bastani, Sören Blomquist and Luca Micheletto, Public Provision of Private Goods, Tagging and Optimal Income Taxation with Heterogeneity in Needs. 47 pp.
- 2010:15 Mattias Nordin, Do voters vote in line with their policy preferences? The role of information. 33 pp.
- 2011:1 Matz Dahlberg, Karin Edmark and Heléne Lundqvist, Ethnic Diversity and Preferences for Redistribution. 43 pp.

- 2011:2 Tax Policy and Employment: How Does the Swedish System Fare? Jukka Pirttilä and Håkan Selin. 73 pp.
- 2011:3 Alessandra Casarico, Luca Micheletto and Alessandro Sommacal, Intergenerational transmission of skills during childhood and optimal public policy. 48 pp.
- 2011:4 Jesper Roine and Daniel Waldenström, On the Role of Capital Gains in Swedish Income Inequality. 32 pp.
- 2011:5 Matz Dahlberg, Eva Mörk and Pilar Sorribas-Navarro, Do Politicians' Preferences Matter for Voters' Voting Decisions?
- 2011:6 Sören Blomquist, Vidar Christiansen and Luca Micheletto, Public provision of private goods, self-selection and income tax avoidance. 27 pp.
- 2011:7 Bas Jacobs, The Marginal Cost of Public Funds is One. 36 pp.