CESifo Working Papers
8575
2020
September 2020

Promoting Education under Distortionary Taxation: Equality of Opportunity versus Welfarism

Pertti Haaparanta, Ravi Kanbur, Tuuli Paukkeri, Jukka Pirttilä, Matti Tuomala

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
https://www.cesifo.org/en/wp
An electronic version of the paper may be downloaded
· from the SSRN website: www.SSRN.com
· from the RePEc website: www.RePEc.org
· from the CESifo website: https://www.cesifo.org/en/wp

CESifo *Working Paper No. 8575*

Promoting Education under Distortionary Taxation: Equality of Opportunity versus Welfarism

Abstract

A common claim in the policy discourse is that a government wishing to achieve equality of opportunity should use public provision of education for equalisation of opportunities rather than income taxation, which only equalizes incomes. We develop a framework in which the tax and education provision rules in the welfarist and non-welfarist/equality of opportunity cases can be transparently compared. We show that in addition to education policies, progressive taxation also plays a role in achieving equality of opportunity, and illustrate how its use may differ under the two objectives. We also show how the provision of public education depends on how private education choices respond, potentially differentially by higher- and lower-income families.

JEL-Codes: H210, H400, O120.

Keywords: educational subsidies, equality of opportunity, income taxation, inequality, public good provision.

Pertti Haaparanta
Aalto University, Espoo / Finland
pertti.haaparanta@aalto.fi

Ravi Kanbur
Cornell University, Ithaca / NY / USA
sk145@cornell.edu

Tuuli Paukkeri
VATT Institute for Economic Research
Helsinki / Funland
tuuli.paukkeri@vatt.fi

Jukka Pirttilä
University of Helsinki, VATT / Finland
jukka.pirttila@helsinki.fi

Matti Tuomala
University of Tampere / Finland
matti.tuomala@tuni.fi

September 10, 2020
This paper is part of UNU-WIDER research programme on ’The Economics and Politics of Taxation and Social Protection’. An earlier version appeared as a WIDER Working Paper. We are grateful for useful comments to seminar audiences at the Nordic Conference for Development Economics, VATT Institute for Economic Research, and the International Institute of Public Finance in 2018. Author declarations of interest: none.

1 Introduction

An often heard refrain in the policy discourse is that rather than use progressive taxation to reduce inequality of incomes, the government should use equal public provision of education to reduce inequality of education, and then let the distribution of income be whatever it turns out to be. Preference for equalizing education over equalizing incomes is sometimes argued for in terms of the presumed greater efficiency, since income taxation would distort the choice between labour effort and leisure. But perhaps a stronger strand in the argument is that equalizing education equalizes opportunities, and that equality of opportunity rather than equality of incomes should be the objective of policy.

Consider, then, an unequal society in which parents spend some of their earned incomes on the education of their children, and this parental input together with equal provision of public education leads to the educational outcomes for children. The government has at its disposal instruments of taxation as well as the level of public provision of education. How should the government choose these instruments in such a setting? The answer depends of course on the government's objectives.

Since the earning of higher incomes requires the use of higher labour effort, the appropriate measure of parental wellbeing is not income per se but utility. One strand of the literature takes as the government's objective a social welfare function defined on the distribution of utilities, which in turn are the outcomes of optimal parental choices on labour, leisure, and expenditure on inputs for children's education. This will be recognized as the classic "welfarist" formulation of the problem emanating from the work of Mirrlees (1971) – welfarist, because the government's objective function depends on, and only on, the "utility outcomes" (of parents in this case).

Contrast this with a "non-welfarist" formulation in which the government cares about, and only about, the distribution of educational outcomes, since this is the distribution of opportunity for the next generation. Parental utility functions do not matter directly in the government's objective function and thus neither do inequalities of utilities or incomes. This follows the arguments of Roemer (1998), who draws on a philosophical tradition going back to Rawls (1971), Dworkin (1981), and Sen (1985), and distinguishes between "circumstances" (factors outside the control of the individual) and "effort" (factors within the individual's control). In this view, inequalities attributable to circumstances are the only legitimate target for government intervention.

This paper frames the difference between "equality of outcomes" and "equality of opportunity" as the distinction between a "welfarist" and a "non-welfarist" objective function. This analytical distinction between welfarist and non-welfarist objective functions makes sharp the informal distinction between "outcomes based" and "opportunities based" objectives in the policy discourse. It allows us to explore in a systematic way the alternative uses of taxation and public education provision under the two types of objectives. Is it the case that under opportunities-based objectives, the tax system used to finance education expenditure should be essentially linear? And if progressive taxation is still used, what does its differential use under the two objectives depend upon? Is it the case that higher provision of equal public education can advance the opportunity based objective? Will the provision of public education in this case necessarily be higher than when the objective is welfarist? These are the types of questions to which the policy discourse gives rise, and that we set out to answer in this paper.

Our paper builds on a large literature on the optimal choice of taxation and public provision of education in the welfarist tradition, including but not limiting to Ulph (1977); Hare and Uplh (1979); Tuomala (1986); Brett and Weymark (2003); Bovenberg and Jacobs (2005); Gasparini and Pinto (2006); Blumkin and Sadka (2008); and Balestrino, Grazzini, and Luporini (2017). In the present paper, much of the analysis is couched in terms of the linear income tax model. The reason is that this tax system – with a lump-sum transfer – is the simplest possible one which is potentially progressive. Therefore, the model is a sufficient framework for examining our main research question, i.e. whether the government wants to tax income in a progressive manner. We also extend the analysis to cover a mixed tax case, a system with non-linear income tax and linear commodity taxes or subsidies.

In addition, our work relates to a recent, growing literature on taxation in an equality of opportunity framework. Most recently, Roemer and Ünveren (2016) set up an intergenerational model in which the current generation makes decisions on education for their children, the future generation. They use public provision of education as the tool to equalize opportunities. The taxes, however, are not used for redistribution but only to finance the public provision of education. Their numerical simulations show that when private acquisition of education is possible, it can undo the intended effect of state provision. The contrast to our paper is that we consider the joint optimisation of taxes as well as public provision of education.

To our knowledge, there is no literature that compares public policies of taxation and education provision by directly comparing the classical welfarist formulation in the tradition of Mirrlees (1971) with the non-welfarist equality of opportunity formulation à la Roemer (1998). Our paper is a first step in this direction. By deriving and presenting optimal taxation and public provision formulae for the two approaches in a comparable manner, we are able to pinpoint the differences between them in a sharp way.

We are also able to place alternative developments in the literature in the context of the contrast between welfarist and non-welfarist frameworks of optimal policy. We contrast our formulation of the objective function for equality of opportunity to the generalized welfare weights approach proposed by Saez and Stantcheva (2016), and argue their model does not fall in a pure "non-welfarist" category. Another related strand of literature is the literature on fair taxation (e.g. Fleurbaey and Maniquet, 2006, 2011). Recent contributions to this literature that are closely related to our paper are Fleurbaey (2006), Valletta (2014), and Fleurbaey and Valletta (2013, 2018). They consider optimal taxation together with goods such as education and health expenditure, which affect the individual's labour productivity and over which they also have direct preferences. We discuss Fleurbaey and Valletta's model extensively and contrast our model to theirs. Even though they extend Valletta's (2014) simpler model by considering a continuum of types and outcomes, and a broader context of human capital investment (which can mean education or health expenditures, or a combination of both), they only consider the case of public subsidies and not of direct public provision. Further, due to multi-dimensional heterogeneity it is quite complicated to obtain more general results from their model. In this paper we present a formulation that relates the Fleurbaey and Valletta formulation to conventional formulations in the literature, allowing for easier comparisons and understandings. We also illustrate how the results differ from the our chosen representation of Equality of Opportunity objectives.

The paper proceeds as follows: Section 2 lays out the basic setup, in which parents with unequal productivities choose labour effort and inputs to children's education to maximize a parental utility function. Section 3 sets out the base results for optimal taxation and public education provision of the welfarist formulation, in which the social welfare function depends only on parental utilities, as the benchmark for later comparison with the equality of opportunity case. Section 4 shows how the optimal tax and public provision formulae are changed when the objective function is non-welfarist, specialized to depending only on the distribution of educational outcomes for children. Sections 2–4 restrict themselves to the case of linear income taxation. Section 5 extends the analysis to non-linear income taxation. In Section 6 we contrast our formulation to two other strands of literature: the generalized welfare weight model by Saez and Stantcheva (2016), and the fairness approach to equality of opportunity in Fleurbaey and Valletta (2018). Section 7 concludes the paper.

2 Individual behaviour

We follow closely the model structure used in Kanbur, Paukkeri, Pirttilä, and Tuomala (2018), allowing for comparison of the results. The framework we have in mind is one where individuals differ in their earning capacity w^i and spend their after-tax income on education and other consumption. The individual budget constraint is $y^i = (1-\tau)z^i + b = x_c^i + x_a^i$, where $z^i = w^i l^i$ denotes labour income, and τ is a linear income tax, which the government uses to finance a lump-sum transfer b. Individual i allocates after-tax income y to private purchases of education, x_c, and other consumption, x_a. Education is thought to benefit the children of the parents who invest in education.

The government can intervene either by public provision of education or by subsidizing private purchases of education. In the first case, utility is $u = u\left[e^i(x_c^i, g), x_a^i, l^i\right]$, where g represents public provision of education. The overall educational level e^i is a function of private purchases and publicly provided education.

When $e^i(x_c^i, g)$ has been fixed by parents and current policies, the welfare of the children is also fixed but for future policies. Assuming that all inherent intertemporal dynamics (like the direct influence of parent's education on child's education) are absent as are also all dynamics in policy making between periods, policies adopted today will be adopted tomorrow. With the assumption that parents ignore the direct impacts of policies (expecting that policies remain unchanged over time) on child's welfare other than those arising directly from education level, our specification of the utility is an approximation of the steady welfare with education levels remaining unchanged between generations. In this case education level can be thought as an indicator of access to welfare. Thereby distribution of education can be thought as an indicator of equality of opportunity. Given this steady state interpretation we do not have to think of which generation is associated with the welfare measure we use as the basis for characterizing optimal policies.

The household maximizes the Lagrangian $u = u\left[e^i(x_c^i, g), x^i, l^i\right] + \lambda\left[(1-\tau)w^i l^i + b - x_c^i - x_a^i\right]$. Its maximum value is denoted by $v^i = u\left[e(x_c^*, g), x_a^*, l^*\right] + \lambda\left[(1-\tau)wl^* + b - x_c^* - x_a^*\right]$. The individual maximization also gives the demand functions $x_c^i = x_c^i(1-\tau, b, g)$ and $x_a^i = x_a^i(1-\tau, b, g)$ as well as labour supply $l^i = l^i(1-\tau, b, g)$.

In the case with no public provision but with a possible educational subsidy s, the budget constraint of the household can be written as $x_a^i + (1-s)x_c^i = (1-\tau)z^i + b$. It is notationally simpler to normalize the situation so that instead of the labour income tax, the government levies consumption taxes on both education and other consumption, and deviations of uniform commodity taxation can be seen as subsidies or taxes on education. Therefore, we work with a model with budget constraint of the form $\sum_j q_j x_j^i = z^i + b$, where $q_j = (p + t_j)$ denotes the consumer price of a good $j = c, a$, with producer prices all equal to p, and t_j represents the tax on good j (a subsidy when $t_j < 0$). Now $v^i(q, b)$ and $x^i(q, b)$ are the indirect utility and consumer demand functions.

3 A welfarist benchmark

3.1 Income taxation

A welfarist government maximizes $\sum_i W\left[v^i(1-\tau, b, g)\right]$ subject to its budget constraint $\sum_i \tau w^i l^i = Nb + N\pi g$, where π is the per-pupil cost of public education and N is the number of households. The first-order conditions, shown in the Appendix, can be used to derive the optimal linear income tax formula:

$$\frac{\tau^*}{1-\tau^*} = \frac{1}{\varepsilon}\left(1 - \frac{z(\beta)}{\bar{z}}\right), \tag{1}$$

where $\beta^i = W' \frac{\partial v^i}{\partial b}$ is the social marginal value of income for person i and $z(\beta) = \frac{\sum \beta^i z^i}{\sum \beta^i}$ denotes the welfare-weighted average income. The elasticity of total income is represented by $\varepsilon = \frac{d\bar{z}}{d(1-\tau)} \frac{(1-\tau)}{\bar{z}}$. The rule is the same as in Kanbur, Paukkeri, Pirttilä, and Tuomala (2018), Section 2.1. The interpretation is the following: when the government has a relatively large welfare weight on the lowest incomes, $z(\beta)$ is small relative to mean income ($\bar{z}$), and the optimal income tax rate is high. On the other hand, the optimal tax rate declines when ε increases.

An alternative way of writing the optimum rule is following Dixit and Sandmo (1977), who utilize the notion of net (of tax revenue) social marginal value of income from Diamond (1975),

$$\rho^i = \frac{\beta^i}{\mu} + \tau w^i \frac{\partial l^i}{\partial b}. \tag{2}$$

Using this definition, the tax rule can be expressed as

$$\tau^* = \frac{-\text{cov}(\rho^i, z^i)}{\frac{1}{N}\sum_i w^i \frac{\partial \tilde{l}^i}{\partial(1-\tau)}}, \tag{3}$$

where $\frac{\partial \tilde{l}^i}{\partial(1-\tau)}$ is the derivative of compensated labour supply. Again, distributional concerns are taken into account in the numerator and the denominator captures efficiency impacts.

3.2 Public provision

When the government directly provides education services, the rule for optimal provision of education is given by

$$\sum_i \beta^i m^i = \mu \left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g}\right), \tag{4}$$

where $m^i = \frac{v^i_g}{\lambda} = \frac{v^i_g}{v^i_b}$ is the marginal rate of substitution for the public good and μ is the Lagrange multiplier of the government budget constraint. This is close to the first-best provision of a publicly provided private good, but the marginal rate of substitution at the left is a weighted one, and at the right a tax revenue term reduces the costs of provision if an increase in public provision increases labour supply.

Following Sandmo (1998), we denote $\gamma = \frac{\mu}{\bar{\beta}}$ and $\bar{\beta} = \frac{1}{N}\sum \beta^i$ and rewrite (4) as

$$N\frac{\sum_i \beta^i m^i}{\sum_i \beta^i} = \gamma\left(N\pi - \sum \tau w^i \frac{\partial l^i}{\partial g}\right). \tag{5}$$

This can be rewritten as

$$\sum_i m^i(1+\delta) = \gamma\left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g}\right), \tag{6}$$

where $\delta = \frac{\text{cov}(\beta^i, m^i)}{\bar{\beta}\bar{m}}$ is the distributional characteristic of publicly provided education, and $\bar{m} = \frac{\sum m^i}{N}$. If the government pays no attention to distributional matters, $\delta = 0$ and the left of Equation (6) is just the conventional sum of the marginal rate of substitution. When distributional concerns matter, the social benefit of public provision increases if the marginal valuation of the publicly provided good is higher for households with low incomes (i.e. high social marginal value of income). In addition, the government needs to take into account the impact of public provision of tax revenues it collects from labour income via the term $\sum_i \tau w^i \frac{\partial l^i}{\partial g}$. If public provision boosts income, then the costs of public provision are reduced relative to the case where public provision would have no impact on tax revenues.

3.3 Commodity taxation: subsidizing education

When the government subsidizes education through commodity taxation, it maximizes $\sum_i W^i\left(v^i(b,q)\right)$ subject to its budget constraint $\sum_i \sum_j t_j x_j^i = Nb$. We assume the government only uses one of the two instruments, so that there is no concurrent public provision of education ($g = 0$).

It is useful to redefine

$$\rho^i = \frac{\beta^i}{\mu} + \sum_j t_j \frac{\partial x_j^i}{\partial b} \tag{7}$$

as the net social marginal utility of income for person i. This notion again takes into account the direct marginal social gain, β^i , and the tax revenue impact arising from commodity demand changes. The rule for optimal commodity taxation for good k is shown to be

$$\sum_i \sum_j t_j \frac{\partial \tilde{x}_k^i}{\partial q_j} = N\text{cov}(\rho^i, x_k^i), \tag{8}$$

where $\tilde{x}_k^i$ refers to compensated demand. The left-hand side of the rule is the aggregate compensated change (weighted by commodity taxes) of good k when commodity prices are changed. The right-hand side refers to the covariance of the net marginal social welfare of income and consumption of the good in question. The rule says that the consumption of those goods whose demand is the greatest for people with low net social marginal value of income (presumably, the rich) should be discouraged by the tax system. Likewise the consumption of goods such as necessities should be encouraged by the tax system. This means that education ought to be subsidized only if its relative valuation is higher among the low-income households.

4 Equality of opportunity

As our framework is strictly paternalistic, we start with a general formulation in which the government maximizes a general paternalistic objective function, $\sum_i P(e^i(x_c^i, g), x_a^i, l^i, g)$. For the general case, the first-order conditions are:

$$\sum_i \frac{dP^i}{d(1-\tau)} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial (1-\tau)} + w^i l^i \right) = 0 \tag{9}$$

$$\sum_i \frac{dP^i}{db} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial b} - 1 \right) = 0 \tag{10}$$

$$\sum_i \frac{dP^i}{dg} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial g} - \pi \right) = 0, \tag{11}$$

where the total derivative is, for example in the case of g, $\frac{dP^i}{dg} = \frac{\partial P^i}{\partial g} + \frac{\partial P^i}{\partial e^i}\frac{\partial e^i}{\partial g} + \frac{\partial P^i}{\partial e^i}\frac{\partial e^i}{\partial x_c^i}\frac{\partial x_c^i}{\partial g} + \frac{\partial P^i}{\partial x_a^i}\frac{\partial x_a^i}{\partial g} + \frac{\partial P^i}{\partial l^i}\frac{\partial l^i}{\partial g}$. In other words, the total impact of extra public provision depends on its direct valuation by the social planner and its indirect impact through consumption and labour supply.

After having derived general tax and public provision rules, we interpret them using societal objectives that only depend on an equitable distribution of education, defining $\sum_i P(e^i, x_a^i, l^i, g) = \sum_i O^i \left\{ e^i \left[x_c^i (1-\tau, b, g), g \right] \right\}$. This function is concave, reflecting inequality aversion in the education space. In this case, $\frac{dP^i}{dg}$ becomes $\frac{dO^i}{dg} = O' \frac{\partial e^i}{\partial g} + O' \frac{\partial e^i}{\partial x_c^i}\frac{\partial x_c^i}{\partial g}$.

4.1 Income taxation

With general paternalistic objectives, the optimal tax rule can be expressed as a combination of a welfarist term and a paternalist corrective term (details in the Appendix):

$$\tau^* = \frac{\text{-cov}(\rho^i, z^i)}{\frac{1}{N}\sum_i w^i \frac{\partial \tilde{l}^i}{\partial(1-\tau)}} + \frac{D}{\frac{1}{N}\sum_i w^i \frac{\partial \tilde{l}^i}{\partial(1-\tau)}}. \tag{12}$$

The first term is the same as in the welfarist case in Equation (3). The second term, where $D = \frac{C_b}{N}\frac{\sum_i z^i}{N} - \frac{C_{(1-\tau)}}{N}$ (within which $C_b = \sum_i \frac{dP^i}{db} - \sum_i \beta^i$, and $C_{1-\tau} = \sum_i \frac{dP^i}{d(1-\tau)} - \sum_i \beta^i z^i$), is a corrective term that takes into account the differences between marginal paternalistic and welfarist valuation of changes in b and $1-\tau$. Due to this term, the tax rate is driven up relative to the welfarist case, if the paternalistic government values the lump-sum benefit b more, or take-home pay $(1-\tau)$ less than the welfarist government. The basic principle that the optimal tax rule is a combination of a welfarist term and a corrective term is in line with the general idea expressed for the non-linear tax by Kanbur, Pirttilä, and Tuomala (2006).

To study the impact of education-only objectives, simply substituting

$\sum_i O^i \left\{ e^i \left[x_c^i \left(1-\tau, b, g\right), g \right] \right\}$ into Equation (12) is not particularly instructive. In the Appendix, we derive the following more intuititve tax rule, which is in line with the welfarist term in Equation (1):

$$\frac{\tau^*}{1-\tau^*} = \frac{1}{\varepsilon}\left(1 - \frac{\tilde{O}}{\bar{z}}\right), \tag{13}$$

where

$$\tilde{O} = \frac{\sum O' \frac{\partial e^i}{\partial x_c^i} \frac{\partial x_c^i}{\partial (1-\tau)}}{\sum O' \frac{\partial e^i}{\partial x_c^i} \frac{\partial x_c^i}{\partial b}}, \tag{14}$$

is the impact of the income tax on education, relative to the effect of income on education. If increasing taxes leads to a large drop in educational attainment, the numerator in Equation (14) goes up, which decreases the tax at the optimum. If, in turn, educational investment becomes more sensitive to income, the denominator in Equation (14) goes up and hence the optimal tax is increased. The higher the income effects—especially at the bottom of the distribution, as they get a higher weight in the social evaluation function—the greater the increase in taxes. A budget-neutral increase in the marginal tax rate also implies a greater lump-sum benefit – that is, a policy that increases progressivity. The implications of this analysis are collected in Proposition 1.

Proposition 1. *A government that only cares about inequality in educational outcomes should also use progressive income taxation to even out educational outcomes, in addition to possibly financing the subsidisation of education. The tax system is more progressive when educational attainment is highly sensitive to income, especially among those at the bottom of* *the educational distribution.*

4.2 Public provision

Consider first a general paternalistic formulation for public provision. It can be written, following Equation (4) as:

$$\sum_i \beta^i m^i = \mu \left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g} \right) - C_g, \tag{15}$$

where $C_g = \sum_i \frac{dP^i}{dg} - \sum_i \beta^i m^i$. In other words, the rule again includes a corrective term that compares paternalistic versus welfarist marginal value of an increase in public provision. If the paternalistic valuation exceeds the welfarist one, the term reduces the costs of public provision.

Turning to public provision under Equality of Opportunity objectives, let us denote $O' \frac{\partial e^i}{\partial x_c^i} \frac{\partial x_c^i}{\partial b} = \beta_O^i$, which is the marginal social (gross) value of income for an Equality of Opportunity government. Let $m_O^i = \left(\frac{de^i}{dg}\right) / \left(\frac{de^i}{db}\right)$ denote the efficiency of public provision in increasing education relative to the income effect. Then, Equation (15) can be written as

$$\sum_i \beta_O^i m_O^i = \mu \left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g} \right), \tag{16}$$

which also implies

$$\sum_i m_O^i (1 + \delta_O) = \gamma_O \left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g} \right), \tag{17}$$

where $\gamma_O = \frac{\mu}{\bar{\beta}_O}$ and $\delta_O = \frac{\mathrm{cov}(\beta_O^i, m_O^i)}{\bar{\beta}_O \bar{m}_O}$ is the distributional characteristic in the Equality of Opportunity case. To interpret the provision rule in Equation (17), notice first that in the case where the distributional characteristic δ_O is zero, implying that the government is not at all averse to inequality in educational attainment, the left-hand side becomes $\sum_i m_O^i$, which captures the relative benefit of affecting the overall educational level via the publicly provided good versus leaving the income to the households. This benefit needs to be weighed against the cost of provision, captured by the first term at the right, $\gamma_O N\pi$ ($= \mu N\pi$ in the case with no distributional concerns). As in the welfarist case, the cost of provision is reduced if the publicly provided good leads to an increase in the tax revenue, captured by the second term (this happens if $\frac{\partial l}{\partial g}$ is positive).

Consider now the influence of aversion against inequality in educational attainment, captured by δ_O. The denominator in m^i_O, $\frac{de^i}{db}$, is always positive, as education is a normal good. Its magnitude can of course vary across individuals. The sign of the numerator in m^i_O, $\frac{de^i}{dg}$, depends on the net impact of public provision on education. As we discussed above, it is likely to be positive, but if public provision is a substitute for private purchases of education at the lower end of the income distribution and a complement at the upper end, the net impact of public provision could well be higher in the upper end. With no distributional concerns, this would increase the benefits of public provision. However, since β^i_O is small for households with high incomes, the covariance in this case would be negative, meaning that education should be under-provided relative to the case with no distributional concerns. Naturally, in the case that m^i_O were higher for households with low incomes, the covariance would become positive, leading to over-provision of education. This discussion is summarized below.

Proposition 2. *Optimal public provision of education for a government whose social welfare function is motivated by Equality of Opportunity concerns is increasing in the impact of public provision on educational attainment relative to the impact of income on education. The provision rule suggests distorting the public provision upwards if the education level is more sensitive to public provision at the lower end of the distribution.*

Note that since also the usefulness of progressive taxation depends on the income effects in educational attainment (Proposition 1), the two instruments are substitutes. Whether tax revenue is used to finance larger lump-sum benefits (more tax progression) or more public education, depends crucially on m^i_O and its distrubution across the population. If income effects are low for low-income individuals and public provision effectively affects educational attainment ($de/db < de/dg$), the government is better off financing education provision directly rather than redistributing taxed income via transfers. Whereas if income effects are relatively high for that group ($de/db > de/dg$), the government should rather rely on income transfers in equalising educational outcomes. For mid-range values of m^i_O, both instruments would play an important role. Rather than posing a strong dichotomy on tax progressivity and public provision of education, the relationship is more subtle and context-specific.

They key issue is hence whether low-income families substitute or complement education by public provision. Peltzman (1973) suggested that public education could crowd out private purchases of schooling, and could even reduce overall schooling consumption. Empirical research has since found some support for this hypothesis for example in the context of U.S. public colleges, though the overall evidence is mixed. (e.g. Castleman and Long 2013; Cellini 2009; Cohodes and Goodman 2014; Long 2004) Slightly more positive results have been found in the context of preschool programmes. Several papers have found the net impact of public provision to be positive, as private provision is either not substituted for public provision, or at least is substituted only partly (e.g. Brinkman, Hasan, Jung, Kinnell, and Pradhan 2017; Bastos and Straume 2016; Bassok, Fitzpatrick, and Loeb 2014; Cascio and Schanzenbach 2013; Cascio 2009).

Only few papers look at heterogeneity of crowding out across income levels. Cohodes and Goodman (2014) find that public college subsidies increased enrollment among the poorest students, even though on net the programme reduced education consumption (as the poorest students formed a small share of the target population). However, Long (2004) finds the opposite, that the poorest students are more sensitive to public subsidies and education crowding out is therefore more severe at the lower end of the income distribution. In the preschool context, Brinkman, Hasan, Jung, Kinnell, and Pradhan (2017) find no heterogeneity between poorer and less poor families in Indonesia, but Cascio and Schanzenbach (2013) find that crowding out is focused among higher-income families in the U.S., as they substitute private care for less expensive public care.

There is not much literature on the income effect on education, but Long's (2004) simulations suggest that changing the in-kind tuition subsidy to public schools into a non-tied grant that can be used in any college, students would consume more education by choosing four-year colleges over two-year colleges, and more selective private colleges over public colleges. Low-income students would be more sensitive to the change than high-income students.

Given the mixed results in the empirical literature, the sign of the numerator of m_O^i is likely to be very context-specific, although we consider it plausible that it would be more positive or less negative for poorer families. There is suggestive evidence that the denominator would be positive, and more strongly so for the disadvantaged students.

4.3 Commodity taxation: subsidizing education

As we are assuming $g = 0$ when the government employs commodity taxation to subsidize education, the objective function $\sum_i O^i \left\{e^i \left[x_c^i(q_c, b, g), g\right]\right\}$ can be simplified as $\sum_i O^i \left(x_c^i(q_c, b)\right)$. This is maximized subject to the budget constraint $\sum_i \sum_j t_j x_j^i - Nb = 0$. It is shown in the Appendix that optimal commodity taxation can be characterized with the rule below:

$$\sum_i \sum_j t_j \frac{\partial \tilde{x}_j^i}{\partial q_k} = N\mathrm{cov}(\rho_O^i, x_k^i) - \frac{1}{\mu} \sum_i O' \frac{\partial \tilde{x}_c^i}{\partial q_k}, \tag{18}$$

where

$$\rho_O^i = O' \frac{\partial x_c^i}{\partial b} \frac{1}{\mu} + \sum_j t_j \frac{\partial x_j^i}{\partial b} \tag{19}$$

is the net social marginal value of income. The left-hand side of the formula above is the compensated aggregated change in the demand of each good. Again, the right-hand side includes a corrective term in comparison to the welfarist rule in Equation (8). Moreover, the covariance rule now measures the relation between the paternalistic net social marginal value of income and the demand for a particular good. According to the second term at the right, when considering the price of education ($k = c$), the demand for education should be encouraged by the tax system, since the own price effect on compensated demand is always negative. This term works towards subsidizing the purchase of education by the tax system. The first term, the covariance term, takes into account distributional concerns, now measured in terms of equality in access to education. If education is highly appreciated by households with high marginal social net value of income (low-income households), this term works towards further effective subsidies on education. In cases where education is valued more by households with low social weight, the covariance term is negative, and it tends to reduce educational subsidies. This leads to the following Proposition:

Proposition 3. *Without direct public provision of education, the desire to equalize educational outcomes works towards encouraging educational attainment by the tax system. The greater the relative price sensitivity of educational attainment among households with higher income, the lower the degree of encouragement through taxes or subsidies.*

5 Non-linear income taxation

5.1 Mixed taxation

We next extend the model where the government can tax or subsidize commodities using linear instruments to cover the case where it can tax income in a non-linear fashion. Income after direct taxation is $y^i = z^i - T(z^i)$, where T denotes any non-linear function. Again, y^i is spent on consumption goods, subject to linear taxes, such that $y^i = \sum_k q_k x_k^i$,where $q_k = p_k + t_k$, with p denoting producer prices. In what follows, we use vector notation and write the budget constraint as $y^i = qx^i$.

It will be useful to utilize the dual approach for this analysis, as in Tuomala (1990). We denote the expenditure function as $E(q,z,w,v)$, which is defined as the minimum expenditure to reach utility $u(x,z,w) = v$. The partially indirect utility is $v(q,b,z,w)$, which results from the household choosing consumption optimally given a budget constraint $qx = b$, where $b = E$ is the expenditure available for the linearly taxed good.

As always in a non-linear income tax problem, we need to take into account the household incentive compatibility constraint. Using the expenditure function, it can be stated as (for any w, w'):

$$E\left[q,z(w),u(x(w),z(w),w),w\right] \le E\left[q,z(w),u(x(w),z(w),w'),w'\right], \tag{20}$$

since the right-hand side is greater than or equal to $qx(w)$. On the other hand, the latter is the same as the left-hand side. This means that $w' = w$ is the value that minimizes the expression at the right. The derivative with respect to w' vanishes at w so that $E_v u_w + E_w = 0$. This serves as the incentive compatibility constraint. Alternatively, it can also be written as

$$v'(w) + \frac{E_w}{E_v} = 0, \tag{21}$$

because $u_w = v'(w)$ by the envelope theorem.

The resource constraint is

$$\int (z - p\tilde{x}) f dw = 0, \tag{22}$$

where $\tilde{x}(q,z,v,w)$ denotes the compensated demand for goods.

Kanbur, Pirttilä, and Tuomala (2006) study in this setting optimal taxation when the government minimizes income poverty, whereas in the present paper the government objective is to achieve a suitable distribution of education, $\int O\left[\tilde{x}_c(q,z,v,w)\right] f dw$. Note that the government objective only depends on one of the consumption goods–private purchases of education–and it is written in terms of compensated demand similarly to the rest of the analysis that follows.

As shown in the Appendix, the rule for optimal commodity taxation can be written as

$$\int t \frac{\partial \tilde{x}}{\partial q} f dw = -\int \omega \frac{\partial x}{\partial w} dw - \int \frac{1}{\mu} O' \frac{\partial \tilde{x}_c}{\partial q} f dw, \tag{23}$$

where $\omega = E_v^{-1} \frac{\alpha}{\mu} > 0$ and α is the Lagrange multiplier of the incentive compatibility constraint, Equation (21). In this formula, the left-hand side is the compensated aggregated change in educational purchases and the first term at the right is the conventional welfarist term. Originally derived by Mirrlees (1976), it states that the consumption of goods that are valued relatively highly by high-ability types, i.e. if $\frac{\partial x}{\partial w} > 0$, should be discouraged by the tax system. Further analysis has shown that this term vanishes if utility is separable between commodity demand and leisure (Atkinson and Stiglitz, 1976).

In addition, there is a new term that measures the impact of commodity taxes on educational purchases. The own price effect is negative, implying that the term is on the whole positive. This works towards encouraging the consumption of education, and this term becomes greater with higher social welfare weight for the household in question (i.e. for low-skilled households) and with more price-elastic demand. This result is summarized below.

Proposition 4. *In an optimal mixed tax system, the consumption of educational services should be encouraged by the tax system. The larger the compensated own-price elasticity of demand, in particular among low-skilled households, the greater the degree of encouragement.*

A corollary to this finding is that even if preferences are separable between commodity demand and leisure, uniform commodity taxation is not optimal. The reason is that the social planner still wants to encourage the consumption of educational services.

We now turn to examining the non-linear part of taxation. For that purpose, one takes the derivative of the Lagrangian in Equation (A.17) with respect to z. The optimality condition is

$$O' \left(\frac{\partial \tilde{x}_c(q,z,v,w)}{\partial z} \right) f - \mu \frac{\partial \tilde{x}}{\partial z} f + \alpha \frac{\partial \left(E_w/E_v\right)}{\partial z} = 0. \tag{24}$$

This expression can be modified (see the Appendix) to obtain a condition for the effective marginal tax rate (i.e. the increase in labour income and commodity taxes when income increases) – which is just the marginal income tax rate in the case where there are no commodity taxes or subsidies:

$$\left(1-t\frac{\partial x}{\partial b}\right)s+1+t\frac{\partial x}{\partial z}=-\frac{1}{f}\omega s_w-\frac{1}{\mu}O'\frac{\partial \tilde{x}_c}{\partial z}, \tag{25}$$

where $s=\frac{v_z}{v_b}$ is the marginal rate of substitution between $z=wl$ and expenditure on goods, $\omega>0$, and s_w is the derivative of the marginal rate of substitution with respect to the ability level. The left-hand side measures the effective marginal tax rate. The first term at the right is the same as in the standard Mirrlees (1976) welfaristic model. The second term at the right is the impact of the Equality of Opportunity concerns on the marginal tax rate. In general it means that the marginal tax rate is not zero at the end points. The last term consists of two components, the first capturing the concavity of the social objective function and the second the link between labour supply and private educational purchases. If an increase in earnings leads to an increase in educational purchases by the households, the last term is on the whole negative, and implies a reduction in the tax rate. The impact of this concern is greater for low-income households, as the social marginal welfare weight tends to be larger for them. These observations lend themselves to the following Proposition:

Proposition 5. *In an optimal mixed tax system, the effective marginal tax rates at the end points are not zero. The effective marginal tax is, ceteris paribus, smaller when labour income and educational purchases are complements. The higher the social marginal value of education at that ability level, the larger is this effect.*

Naturally, when income and education are substitutes, effective marginal tax rates tend to increase. These mechanisms serve as a way for the government to indirectly influence the educational level via labour supply. An interesting case is one in which income increases at one ability level raise the demand for education and lead to reductions in education at another. Consider, for instance, a situation in which at low ability levels income and education are complements, whereas they would be substitutes at higher ability levels. This would mean that the effective marginal tax rate tends to go down at low income levels and is pushed upwards at higher incomes. As always in optimal tax research, one needs to remember that this reasoning is only valid when other things are equal, and these other things may not remain intact as the optimality conditions are evaluated at different levels when making comparisons between traditional welfarist versus non-welfarist analyses.

5.2 Public provision

We finally consider the case of public provision. For brevity, subsidies and other indirect taxes are assumed away, but enlarging the analysis to cover them would be straightforward along the lines of the analysis in Pirttilä and Tuomala (2004, Section 5). The government objective function is now written as $\int O\left[e\left(\tilde{x}_c(z,v,w,g),g\right)\right] f dw$. Here, we use the dual approach and work with the expenditure function x refers to the minimum expenditure required to reach utility v.

We show in the Appendix how we arrive at the following provision rule:

$$\int \pi f dw = \int m f dw - \int \omega m_w dw + \int \frac{1}{\mu} O' \left(\frac{\partial e}{\partial x_c} \frac{\partial \tilde{x}_c(z,v,w,g)}{\partial g} + \frac{\partial e}{\partial g} \right), \qquad (26)$$

where $m = \frac{v_g}{v_b}$ and m_w its derivative with respect to the wage rate. The rule compares the marginal cost of public provision (the left-hand side) with the marginal benefits (right-hand side). The first two terms are familiar from the welfarist literature. They measure the willingness to pay for the public provision and the way this willingness is linked with the ability level. The last term at the right is novel: it measures the impact of public provision on equality of education. The greater the overall impact, both directly and indirectly via private purchases, the higher the marginal benefits of public provision. Since the function O is concave, the rule is more sensitive to how public provision changes education at the lower end of the educational distribution. Again the provision rule can be expressed as a sum of a welfarist term and an equality of opportunity corrective term.

6 Alternative approaches

In this section, we connect the considerations reviewed above to two further approaches to Equality of Opportunity.

6.1 Interpretation using generalized social marginal welfare weights

Saez and Stantcheva (2016) propose a framework of generalized marginal social welfare weights, as well as an extension of the framework into modeling Equality of Opportunity. These weights are represented by $\xi^i(x_c^i, x_a^i, z^i, \chi^{i,u}, \chi^{i,b}, \chi^{i,s})$. Here, $\chi^{i,u}$ denotes characteristics that enter the private utility function, $\chi^{i,s}$ those that are accounted for only by the social planner, and $\chi^{i,b}$ those characteristics that affect both individual and social welfare. As Saez and Stantcheva (2016) only illustrate their approach in the case of income tax alone, we extend it to cover public provision of education and subsidisation through commodity taxation. As they show in their online appendix, in the case where the individual utility is a money-metric one, the approach can be thought of as if the government were maximizing $\sum_i \xi^i v^i$. When indirect utility is money-metric, the social marginal value of income to individual i is just ξ^i. If the government were welfaristic with a social welfare function of $W\{v\}$, then $\xi^i = \frac{\partial W}{\partial v^i}$.

Identically to the derivation in Section 3.2, the public good provision rule then simply becomes

$$\sum_i m^i(1+\delta_{SS}^i) = \gamma \left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g} \right), \tag{27}$$

where $\delta_{SS}^i = \frac{\mathrm{cov}(\xi^i, m^i)}{\bar{\xi}\bar{m}}$ is the distributional characteristic of publicly provided education, now defined on the basis of ξ.

In the case of commodity taxation, the optimal tax rule is of the form

$$\frac{1}{N}\sum_i \sum_j t_j \frac{\partial \tilde{x}_k^i}{\partial q_j} = \mathrm{cov}(\rho_{SS}^i, x_k^i), \tag{28}$$

where $\rho_{SS}^i = \frac{\xi^i}{\mu} + \sum_j t_j \frac{\partial x_j^i}{\partial b}$.

The Saez–Stantcheva (SS) approach thus produces similar public provision and education subsidisation rules as the standard welfarist model. However, it only works for such social preferences that are not paternalistic – that is, they accept individual welfare as a starting point. Therefore, our formulation above, where $O(e)$ is a function of education alone and does not put any welfare weight to the consumption of other goods or leisure, is not compatible with the SS approach.

6.2 Equality of Opportunity as Fairness

The last approach to equality of opportunity we apply is a version of the fairness theory developed by, for example, Fleurbaey (2008) and Fleurbaey and Maniquet (2011). It is closely related to the theory of equality of opportunity, as the fairness theory seeks a balance between reward (right to fruits of own effort) and compensation (right for compensation due to bad circumstances beyond individual control). The questions studied in this paper have been studied from the fairness point of view in a closely related paper by Fleurbaey and Valletta (2018) focusing on optimal non-linear income taxation. We use and extend the linear taxation version presented in the working paper version of the article (Fleurbaey and Valletta, 2013). In this section, we i) show that the results from the Fleurbaey–Valletta model of fairness can be formally presented in a way similar to results in the previous sections, improving their comparability; ii) show that the results from the Fleurbaey–Valletta model and from the models used in the previous sections are closely related, but not identical, for linear income taxation and education subsidies; and iii) show more detailed characterizations of the optimal fair policies and extend the fair tax model with commodity taxation and public provision of education.

One important difference is that in the Fleurbaey–Valletta model, education improves personal productivity, instead of increasing individual welfare as in the other models used in this paper. But in both approaches, the key is the education production function: that the education level is a function of private investment and public provision of education. Fleurbaey and Valletta model this as an individual cost of obtaining a certain level of education for a given level of public provision. The cost function is taken as a circumstance facing individuals, and hence not their responsibility. We can implement this approach by taking the education production function used above, $e^i\left(x_c^i, g\right)$, and inverting it to find the cost of obtaining a given level of education:

$$x_c^i = x_c^i\left(e^i, g\right), \frac{\partial x_c^i}{\partial e^i} > 0, \frac{\partial x_c^i}{\partial g} \leq 0. \tag{29}$$

The money-metric welfare is obtained by asking: with everybody facing the same circumstances, what lump-sum income transfer would make an individual indifferent between her present state and the state in which she faces the equalized circumstances? Fleurbaey and Valletta argue that the relevant circumstances are the average productivity and the average cost of education. Thus the transfer needed to make the individual indifferent between her present state and the state with harmonized circumstances is the value function of the optimization problem:

$$\begin{aligned} \min \quad & x_a + \bar{x}_c(e,g) - \bar{w}l \\ s.t. \quad & u^i(e,x,l) \geq u^i\left(e^i, x^i, l^i\right). \end{aligned} \tag{30}$$

Here, $u^i\left(e^i,x^i,l^i\right)$ is the welfare of individual i at the current allocation of resources. Thus, the value function (the transfer) for individual i is $\vartheta^i = \vartheta^i\left(\bar{x}_c, \bar{w}, u^i\left(e^i,x^i,l^i\right)\right)$.

Note that we allow for heterogeneity in individual utility functions and assume individuals to be responsible for their preferences. An individual's welfare is, for the case of linear income tax and public provision of education, given by the indirect utility function $v^i(t,b,g)$, as above, and is analogous for the case of commodity taxes. Social welfare is maximized by maximizing the welfare of the worst-off person.[1] We give this person index o (there are N individuals in total).

Our main results specify exact conditions that the worst-off person's consumption patterns, willingness to pay, and investments have to hold for commodity taxation to favour the worst-off person, the social cost of public provision to be reduced, and private investment to education to be taxed or subsidized. These are more detailed than obtained in the welfarist or Equality of Opportunity approach analysed above. The details of the derivation are presented in the Appendix.

6.2.1 Linear taxation

Optimal policies maximize the money-metric measure of the worst-off person,

$\vartheta^o\left(\bar{w}, \bar{c}, v^i(1-\tau, b)\right)$. In the Appendix we show that the tax rule satisfies

$$\frac{\tau^*}{1-\tau^*} = \frac{1}{\sum_i \theta^i \varepsilon^i_{l,1-\tau}}\left(1 - A\theta^o\right). \tag{31}$$

Here, $A \equiv 1 - \tau \sum_i \frac{w^i l^i}{Nb} \varepsilon^i_{l,b} > 0$, and $\varepsilon^i_{y,x}$ is the elasticity of y with respect to x. θ^i denotes the

[1] Fleurbaey and Valletta (2018) discuss conditions for the existence of a worst-off person. Note also that this is not, in general, equivalent to calculating the maximin policies in a welfarist setting. Here optimization is based on money-equivalent measures of welfare which are affected by the salient circumstances. As pointed out above, the fairness approach is not a special case of welfarism.

share of individual i's income in total income, $\theta^i \equiv \frac{w^i l^i}{\sum_i w^i l^i}$. The tax rate is positive and below unity as long as $\sum_i \theta^i \varepsilon^i_{l,1-\tau} > 0$, which is plausible, and when $A\theta^o < 1$. The formula in Equation (31) is analogous to our results for linear tax in the other cases. The difference is that it focuses on the income of the worst-off citizen relative to the average income as the key parameter. In other words, the theory proposes this ratio as the key parameter for analysing the fairness of linear income tax systems.[2]

6.2.2 Public provision

The optimality conditions in the fairness case can be written in exactly the same format as in the welfarist case above. This is as the fair social welfare function gives $\beta^o_F \equiv \frac{\partial \vartheta^o}{\partial v^o}\frac{\partial v^o}{\partial b}$ as the marginal social welfare weight of the worst-off person, while the weight for the others is $\beta^i_F = 0 \forall i \neq o$ (as $\frac{\partial \vartheta^o}{\partial v^i} = 0$). Thus, the public provision rule can be expressed as follows:

$$\beta^o_F m^o = \mu \left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g} \right). \tag{32}$$

This is the fairness equivalent to the welfarist public provision rule in Equation (4) and Equality of Opportunity rule in (16). However, it is difficult to infer from Equation (32) what it implies for public education compared to welfarist public provision or to Samuelson-efficient provision. To get ahead, we use again $\beta^o_F m^o = \sum_i \beta^i_F m^i$, and rewrite Equation (32) as, equivalent to Equations (6) and (17):

$$\sum_i m^i (1 + \delta_F) = \mu \left(N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g} \right), \tag{33}$$

where $\delta_F = \frac{\operatorname{cov}(\beta^i_F, m^i)}{\bar{\beta}_F \bar{m}}$, with $\bar{\beta}_F \equiv \frac{\sum_i \beta^i_F}{N} = \frac{\beta^o_F}{N}$.

The "fair" demand for public education is higher (or the cost of public provision lower) than proposed by the Samuelson rule if $\delta_F > 0$. This holds if and only if

$$m^o > \frac{\sum_{i=1}^{N-1} m^i}{N-1}, \tag{34}$$

otherwise the demand is reduced. Thus, if the worst-off person values education more than the other citizens on average, the fairness criterion suggests, ceteris paribus, extension of public

[2] Note that Fleurbaey and Valletta (2013) do not express the optimal income tax formula in a similar way as here. The formulation here focuses directly on income distribution statistic as a determinant of optimal tax as in the other approaches considered.

education. This is of course not necessarily the case.

Note the differences to the welfarist and Equality of Opportunity cases (Equations (6) and (17)). In both cases, distributional concerns play an important role. For the welfarist case, if the social value of income covaries positively with the private valuation of education, it increases the value of public provision. For an Equality of Opportunity minded government, the distribution of the relative effectiveness of public provision relative to income in increasing the educational level is what matters. In the fairness case, only the private valuation of education, and only by the worst-off person matters.

The optimal education policy can also be expressed so as to highlight the importance of the willingness to pay for public education by the worst-off person, m^o:

$$m^o = \frac{N\pi - \sum_i \tau w^i \frac{\partial l^i}{\partial g}}{N - \sum_i \tau w^i \frac{\partial l^i}{\partial b}}. \tag{35}$$

This can be rewritten as

$$Nm^o = \frac{N\pi - \sum_i \tau \frac{w^i l^i}{g} \varepsilon^i_{l,g}}{1 - \sum_i \tau \frac{w^i l^i}{Nb} \varepsilon^i_{l,b}}. \tag{36}$$

Here, $\varepsilon^i_{l,b} \equiv \frac{\partial l^i}{\partial b} \frac{b}{l^i}$ is the income elasticity of labour supply and $\varepsilon^i_{l,g}$ is the corresponding elasticity with respect to public provision of education. We know that $\varepsilon^i_{l,b} < 0$ if leisure is a normal good, as we assume.

This formulation is the equivalent to the standard optimality condition for optimal public good production. The left-hand side of Equation (36) gives the willingness to pay for public education. As Fleurbaey and Valletta (2018) argue, the willingness to pay for public education by the worst-off person is important because the worst-off person is not only a deserving poor but also a person with a high cost of education (e.g. due to the high cost of reaching a given level of education). This can arise, for example, if public education is of low quality and therefore the education level is very insensitive to changes in public education: $\frac{\partial e^o}{\partial g}$ is small. In that case public provision does not reduce the overall cost of education, as private investment is still needed: also $\frac{\partial x^i_c}{\partial g}$ in the education cost function is small. This would mean that public education is not very effective in improving the social welfare (nor the welfare of the worst-off person). If this were the case, it would raise the possibility that taxing education is optimal.

6.2.3 Commodity taxation

Using β_F^i as above ($\beta_F^o \equiv \frac{\partial \vartheta^o}{\partial v^o}\frac{\partial v^o}{\partial b}$; $\beta_F^i = 0 \forall i \neq o$), the optimal commodity taxes satisfy the condition, similar to the welfarist case in Equation (8):

$$\sum_i \sum_j t_j \frac{\partial \tilde{x}_k^i}{\partial q_j} = N\mathrm{cov}\left(\rho_F^i, x_k^i\right), \tag{37}$$

where

$$\rho_F^i = \frac{\beta_F^i}{\mu} + \sum_j t_j \frac{\partial x_j^i}{\partial b} \tag{38}$$

is the net social marginal value of income in the fairness case. As in the welfarist case, this means that if the individuals with the greatest social weight tend consume less of good k than the people with less weight, the commodity taxes are set to reduce the consumption of the good. In the fairness case we can actually say something more specific. It can be shown that the covariance is negative if and only if:

$$x_k^o < \sum_{i=1}^{N-1} \frac{\alpha^i}{\sum_i \alpha^i} x_k^i, \tag{39}$$

where $\alpha^i \equiv 1 - \sum_j t_j \frac{\partial x_j^i}{\partial b}$. The consumption of good k by the worst-off person must be below a weighted average of the consumption of the same good by other individuals for the commodity taxes to punish the consumption of the good. Consumption of goods relatively heavily used by the worst-off person should be subsidized.

The results are analogous to the results in Fleurbaey and Valletta (2013) as well as those in Fleurbaey and Valletta (2018) for the case of non-linear taxation: in the latter it is shown that, in general, it is not optimal to treat education investment as deductible, and education investment should be taxed when the worst-off person does not value it much. Note again the difference to the Equality of Opportunity minded rule (Equation (18)), where there is pressure to encourage the demand for education even when the lower-income individuals value it less than higher-income individuals.

In sum, the optimal fair tax and education rules can be expressed in equivalent form to the welfarist and Equality of Opportunity rules. Whereas the latter two depend clearly on the distribution of the value of education across the population, the fairness approach achieves this by emphasising the worst-off individual's valuation of education compared to the rest of the population. Another key difference stems from the paternalism inherent in the Equality of Opportunity approach: it does not take individual valuations into account, but only the effectiveness of different policies in achieving the paternalistic objective matters.

7 Conclusion

Let us return to the four questions posed in the Introduction, which emerge from the policy discourse. Is it the case that progressive taxation is not used at all under opportunities-based objectives? We have shown that the argument of "progressive taxation for welfarist objectives and equal provision of public education for equality of opportunity objectives" poses a false dichotomy. Progressive taxation is a potent instrument for equalizing opportunity through equalizing education outcomes. What does the differential use of progressive taxation under the two objectives depend upon? We have derived and presented optimal tax formulae in a way that facilitates comparison between the two regimes. When educational outcomes are highly sensitive to parental inputs relative to public provision, perhaps paradoxically the case for progressive taxation tends to be stronger under the equality of opportunity objective.

Is it the case that higher provision of public education can advance the opportunity-based objective? Will the provision of public education in this case necessarily be higher than when the objective is welfarist? We have shown how answers to these questions depend on the nature of the "education production function" – the precise way in which parental and public inputs go together to produce educational outcomes for children. The extent of public provision is relatively low, if education is valued relatively more by high-income households (as might well be the case).

The answers to these questions illustrate how our framework can help to address specific questions in the policy discourse. Our analysis has, however, been wider ranging. We have used our framework to assess commodity taxation, where we get results on education subsidies similar to those on public provision. We have analysed non-linear income taxation under the two regimes and shown that, unlike for the welfarist case, for equality of opportunity the effective marginal tax rates should not be set to zero at the end points of ability distribution. We have also highlighted how the generalized welfare weights framework of Saez and Stantcheva (2016) cannot fully capture the non-welfarism inherent in equality of opportunity objectives.

Equality of opportunity has emerged as a major framework for the public policy discourse. This paper has attempted to present a framework in which the consequences of this framework can be compared to those of the welfarist literature. In the process we have asked and answered a number of specific questions on taxation and public provision to show the utility of the formulation. In particular, we have shown that progressive taxation and equality of opportunity are not opposed to each other. A rich research agenda lies ahead.

References

ATKINSON, A. B., AND J. E. STIGLITZ (1976): "The design of tax structure: direct versus indirect taxation," *Journal of Public Economics*, 6(1-2), 55–75.

BALESTRINO, A., L. GRAZZINI, AND A. LUPORINI (2017): "A normative justification of compulsory education," *Journal of Population Economics*, 30(2), 537–67.

BASSOK, D., M. FITZPATRICK, AND S. LOEB (2014): "Does state preschool crowd-out private provision? The impact of universal preschool on the childcare sector in Oklahoma and Georgia," *Journal of Urban Economics*, 83(September), 18–33.

BASTOS, P., AND O. R. STRAUME (2016): "Preschool education in Brazil: Does public supply crowd out private enrollment?," Policy research working paper 7573, Washington, DC: World Bank.

BLUMKIN, T., AND E. SADKA (2008): "A case for taxing education," *International Tax and Public Finance*, 15(2), 145–63.

BOVENBERG, A. L., AND B. JACOBS (2005): "Redistribution and education subsidies are Siamese twins," *Journal of Public Economics*, 89(11-12), 2005–35.

BRETT, C., AND J. A. WEYMARK (2003): "Financing education using optimal redistributive taxation," *Journal of Public Economics*, 87(11), 2549–69.

BRINKMAN, S. A., A. HASAN, H. JUNG, A. KINNELL, AND M. PRADHAN (2017): "The impact of expanding access to early childhood education services in rural Indonesia," *Journal of Labor Economics*, 35(S1), S305–35.

CASCIO, E. U. (2009): "Maternal labor supply and the introduction of kindergartens into American public schools," *The Journal of Human Resources*, 44(1), 140–70.

CASCIO, E. U., AND D. SCHANZENBACH (2013): "The impacts of expanding access to high-quality preschool education," Nber working paper 19735, Cambridge: National Bureau of Economic Research.

CASTLEMAN, B. L., AND B. T. LONG (2013): "Looking beyond enrollment: The causal effect of need-based grants on college access, persistence, and graduation," *Journal of Labor Economics*, 34(4), 1023–73.

CELLINI, S. R. (2009): "Crowded colleges and college crowd-out: The impact of public subsidies on the two-year college market," *American Economic Journal: Economic Policy*, 1(2), 1–30.

COHODES, S. R., AND J. S. GOODMAN (2014): "Merit aid, college quality, and college completion: Massachusetts' Adams Scholarship as an in-kind subsidy," *American Economic Journal: Applied Economics*, 6(4), 251–85.

DIAMOND, P. (1975): "A many-person Ramsey tax rule," *Journal of Public Economics*, 4(4), 335–42.

DIXIT, A., AND A. SANDMO (1977): "Some simplified formulae for optimal income taxation," *Scandinavian Journal of Economics*, 79(4), 417–23.

DWORKIN, R. (1981): "What is equality? Part 2: Equality of resources," *Philosophy and Public Policy*, 4(10), 283–345.

FLEURBAEY, M. (2006): "Is commodity taxation unfair?," *Journal of Public Economics*, 90(10-11), 1765–87.

——— (2008): *Fairness, responsibility, and welfare*. Oxford: Oxford University Press.

FLEURBAEY, M., AND F. MANIQUET (2006): "Fair income tax," *Review of Economic Studies*, 73(1), 55–83.

——— (2011): *A theory of fairness and social welfare*. Cambridge: Cambridge University Press.

FLEURBAEY, M., AND G. VALLETTA (2013): "Fair optimal tax with endogenous productivities," Research memorandum 024, Maastricht: Maastricht University, Graduate School of Business and Economics (GSBE).

——— (2018): "Fair optimal tax with endogenous productivities," *Journal of Public Economic Theory*, 20(6), 849–73.

GASPARINI, L. C., AND S. M. PINTO (2006): "Equality of opportunity and optimal cash and in-kind policies," *Journal of Public Economics*, 90(1), 143–69.

HARE, P., AND D. UPLH (1979): "On education and distribution," *Journal of Political Economy*, 87(5), S193–S212.

KANBUR, R., T. PAUKKERI, J. PIRTTILÄ, AND M. TUOMALA (2018): "Optimal taxation and public provision for poverty reduction," *International Tax and Public Finance*, 25(1), 64–98.

KANBUR, R., J. PIRTTILÄ, AND M. TUOMALA (2006): "Non-welfarist optimal taxation and behavioural public economics," *Journal of Economic Surveys*, 20(5), 849–68.

LONG, B. T. (2004): "Does the format of a financial aid program matter? The effect of state in-kind tuition subsidies," *Review of Economics and Statistics*, 86(3), 767–82.

MIRRLEES, J. A. (1971): "An exploration in the theory of optimum income taxation," *Review of Economic Studies*, 38(114), 175–208.

MIRRLEES, J. A. (1976): "Optimal tax theory: A synthesis," *Journal of Public Economics*, 6(4), 327–58.

PELTZMAN, S. (1973): "The effect of government subsidies-in-kind on private expenditures: The case of higher education," *Journal of Political Economy*, 81(1), 1–27.

PIRTTILÄ, J., AND M. TUOMALA (2004): “Poverty alleviation and tax policy,” *European Economic Review*, 48(5), 1075–90.

RAWLS, J. (1971): *A theory of justice*. Cambridge: Harvard University Press.

ROEMER, J. E. (1998): *Equality of opportunity*. Cambridge: Harvard University Press.

ROEMER, J. E., AND B. ÜNVEREN (2016): “Dynamic equality of opportunity,” *Economica*, 84(334), 322–43.

SAEZ, E., AND S. STANTCHEVA (2016): “Generalized social welfare weights for optimal tax theory,” *American Economic Review*, 106(1), 24–45.

SANDMO, A. (1998): “Redistribution and the marginal cost of public funds,” *Journal of Public Economics*, 70(3), 365–82.

SEN, A. (1985): *Commodities and capabilities*. Amsterdam: Elsevier, North Holland.

TUOMALA, M. (1986): “On the optimal income taxation and educational decisions,” *Journal of Public Economics*, 30(2), 183–98.

——— (1990): *Optimal income tax and redistribution*. Oxford: Oxford University Press.

ULPH, D. (1977): “On the optimal distribution of income and educational expenditure,” *Journal of Public Economics*, 8(3), 341–56.

VALLETTA, G. (2014): “Health, fairness and taxation,” *Social Choice and Welfare*, 43(1), 101–40.

A Appendix: Proofs

A.1 Welfarist benchmark

A.1.1 Income taxation

The Lagrangian for a welfarist government is $\sum_i W\left[v^i(1-\tau,b,g)\right]+\mu\left(\sum_i \tau w^i l^i - Nb - N\pi g\right)$. The first-order conditions of the government optimization problem with respect to $1-\tau$, b, and g, are:

$$\sum_i W' \frac{\partial v^i}{\partial(1-\tau)} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial(1-\tau)} - w^i l^i\right) = 0 \tag{A.1}$$

$$\sum_i W' \frac{\partial v^i}{\partial b} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial b} - 1\right) = 0 \tag{A.2}$$

$$\sum_i W' \frac{\partial v^i}{\partial g} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial g} - \pi\right) = 0 \tag{A.3}$$

Equations (A.1) and (A.2) can be used to derive the optimal linear income tax in Equation (1) (see Kanbur, Paukkeri, Pirttilä, and Tuomala (2018)). To see this, multiply (A.2) by $\frac{\sum wl}{N}$ and divide (A.1) by N. Then subtract the former from the latter to get

$$\left\{\sum_i \beta^i w^i l^i \frac{1}{N} - \sum_i \frac{\beta^i}{N} \sum_i \frac{w' l^i}{N}\right\} + \mu\tau\left\{\sum_i w^i \frac{\partial l^i}{\partial(1-\tau)} - \sum_i w^i \frac{\partial l^i}{\partial b} \frac{1}{N} \sum_i \frac{w' l^i}{N}\right\} = 0. \tag{A.4}$$

Using the Slutsky equation $\left(\frac{\partial l^i}{\partial(1-\tau)} = \frac{\partial \tilde{l}^i}{\partial(1-\tau)} + \frac{\partial l^i}{\partial b} wl\right)$, where tilde refers to compensated labour supply, and the notion on net marginal social welfare (the expression in (2)), one can arrive at the formula in (3).

A.1.2 Public provision

Equation (4) follows directly from rearranging (A.3) and using the definition of $m^i = \frac{v^i_g}{\lambda} = \frac{v^i_g}{v^i_b}$.

A.1.3 Commodity taxation

The Lagrangean is $\sum_i W^i\left(v^i(b,q)\right) + \mu\left[\sum_i \sum_j t_j x^i_j - Nb\right]$ and the first-order condition with respect to t_k :

$$-\sum_i \beta x_k + \mu \sum_i \left(x^i_k + \sum_j t_j \frac{\partial x^i_j}{\partial t_k}\right) = 0 \tag{A.5}$$

Using the definition of $\rho^i = \frac{\beta^i}{\mu} + \sum_j t_j \frac{\partial x^i_j}{\partial b}$, the Slutsky equation, and Slutsky symmetry, (A.5) can be rewritten as

$$\sum_i \sum_j t_j \frac{\partial x^i_j}{\partial t_k} = N \left(\sum \frac{1}{N}_i \rho x_k - \sum_i \frac{1}{N} x^i_k \right),$$

from which the rule in (8) follows, using the property that $\sum \frac{1}{N} \rho^i = 1$.

A.2 Equality of Opportunity

A.2.1 Income taxation

Equation (12) can be obtained as follows. First rewrite the first-order conditions (9) and (10) by adding and subtracting terms as:

$$\sum_i \beta^i z^i + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial(1-\tau)} - w^i l^i \right) + \sum_i \frac{dP^i}{d(1-\tau)} - \sum_i \beta^i z^i = 0 \quad \text{(A.6)}$$

$$\sum_i \beta^i + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial b} - 1 \right) + \sum_i \frac{dP^i}{db} - \sum_i \beta^i = 0. \quad \text{(A.7)}$$

Denote $\sum_i \frac{dP^i}{d(1-\tau)} - \sum_i \beta^i z^i = C_{(1-\tau)}$ and $\sum_i \frac{dP^i}{db} - \sum_i \beta^i = C_b$. Multiply Equation (A.7) by $\frac{1}{N} \frac{\sum_i z^i}{N}$ and divide Equation (A.6) by N. Then subtract the former from the latter to get:

$$\frac{\sum_i \beta^i z^i}{N} - \frac{\sum_i \beta^i}{N} \frac{\sum_i z^i}{N} + \frac{\mu \tau}{N} \left(\sum_i w^i \frac{\partial l^i}{\partial(1-\tau)} - \sum_i w^i \frac{\partial l^i}{\partial b} \frac{\sum_i z^i}{N} \right) + \sum_i \frac{dP^i}{d(1-\tau)} + \left(\frac{C_{(1-\tau)}}{N} - \frac{C_b}{N} \frac{\sum_i z}{N} \right) = 0. \quad \text{(A.8)}$$

Collecting terms then yields the result in the main text.

When the social objective is to achieve an equal distribution of education, the first-order conditions governing the choice of the tax rate are:

$$\sum_i O' \frac{\partial e^i}{\partial x^i_c} \frac{\partial x^i_c}{\partial(1-\tau)} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial(1-\tau)} - w^i l^i \right) = 0 \quad \text{(A.9)}$$

$$\sum_i O' \frac{\partial e^i}{\partial x^i_c} \frac{\partial x^i_c}{\partial b} + \mu \sum_i \left(\tau w^i \frac{\partial l^i}{\partial b} - 1 \right) = 0. \quad \text{(A.10)}$$

Dividing these two yields:

$$\frac{\sum_i O' \frac{\partial e^i}{\partial x^i_c} \frac{\partial x^i_c}{\partial(1-\tau)}}{\sum_i O' \frac{\partial e}{\partial x^i_c} \frac{\partial x^i_c}{\partial b}} = \frac{w^i l^i - \tau \sum_i w^i \frac{\partial l^i}{\partial(1-\tau)}}{1 - \tau \sum_i w^i \frac{\partial l^i}{\partial b}}. \quad \text{(A.11)}$$

Following the steps in Kanbur, Paukkeri, Pirttilä, and Tuomala (2018, p. 83-84) yields the rule in Equation (13).

A.2.2 Public provision

The rule in (15) follows from rearranging the first-order condition of g and adding and subtracting $\sum_i \beta^i m^i$.

A.2.3 Commodity taxation

The first-order conditions are:

$$\sum_i O' \frac{\partial x_c^i}{\partial b} + \mu \left(\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b} - N \right) = 0 \tag{A.12}$$

$$\sum_i O' \frac{\partial x_c^i}{\partial q_k} + \mu \left(\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial q_k} + \sum_i x_k^i \right) = 0. \tag{A.13}$$

The first one of these can be used if the government is allowed/able to set the demogrant optimally. Denote again

$$\rho_O^i = O' \frac{\partial x_c^i}{\partial b} \frac{1}{\mu} + \sum_j t_j \frac{\partial x_j^i}{\partial b} \tag{A.14}$$

as the net social marginal utility of income for person i. Equation (A.12) implies that $\bar{\rho}_O = 1$. Using the Slutsky equation $\frac{\partial x_j^i}{\partial q_k} = \frac{\partial \tilde{x}_j^i}{\partial q_k} - x_k^i \frac{\partial x_j^i}{\partial b}$, where $\tilde{x}$ depicts compensated demand, one can rewrite Equation (A.13) as:

$$\sum_i O' \frac{1}{\mu} \left(\frac{\partial \tilde{x}_c^i}{\partial q_k} - x_k^i \frac{\partial x_c^i}{\partial b} \right) + \left[\sum_i \sum_j t_j \left(\frac{\partial \tilde{x}_j^i}{\partial q_k} - x_k^i \frac{\partial x_j^i}{\partial b} \right) + \sum_i x_k^i \right] = 0. \tag{A.15}$$

With the help of Equation (A.14), Slutsky symmetry and by rearrangement, this can be written as

$$\sum_i \sum_j t_j \frac{\partial \tilde{x}_j^i}{\partial q_k} = \sum_i \rho_O^i x_k^i - \sum_i x_k^i - \frac{1}{\mu} \sum_i O' \frac{\partial \tilde{x}^i}{\partial q_k}, \tag{A.16}$$

which can also be expressed in a covariance format (Equation (18)) in the main text.

A.2.4 Mixed taxation

In the mixed tax case, the government maximizes its social welfare function subject to the constraints in Equations (21) and (22). The Lagrangian of this problem is

$$
\begin{aligned}
L &= \int \left\{ \left(O\left[\tilde{x}_c(q,z,v,w)\right] + \mu\left(z - p\tilde{x}\right)\right) f + \alpha v'(w) + \alpha \frac{E_w}{E_v} \right\} dw \\
&= \int \left\{ \left(O\left[\tilde{x}_c(q,z,v,w)\right] + \mu\left(z - p\tilde{x}\right)\right) f - \alpha' v(w) + \alpha \frac{E_w}{E_v} \right\} dw \\
&\quad + \alpha(\infty)v(\infty) - \alpha(0)v(0),
\end{aligned}
\tag{A.17}
$$

where f is the distribution function of abilities and where the equality follows from integration by parts.

Consider first the first-order condition with respect to commodity prices, q, which is given by

$$
\int O' \left(\frac{\partial \tilde{x}_c(q,z,v,w)}{\partial q} \right) f dw - \int \left(\mu p \frac{\partial \tilde{x}(q,z,v,w)}{\partial q} f + \alpha \frac{\partial \left(E_w/E_v\right)}{\partial q} \right) dw = 0. \tag{A.18}
$$

The rule for optimal commodity taxes in the case of mixed taxation can be derived as follows. Note first that because of the properties of the expenditure function,

$$
\begin{aligned}
\frac{\partial \left(E_w/E_v\right)}{\partial q} &= \frac{\left(E_{wq}E_v - E_{vq}E_w\right)}{E_v^2} = \frac{\left(\frac{\partial \tilde{x}}{\partial w} - \left(E_w/E_v\right)\frac{\partial \tilde{x}}{\partial v}\right)}{E_v} \\
&= E_v^{-1}\left(\frac{\partial \tilde{x}}{\partial w} - u_w \frac{\partial \tilde{x}}{\partial v}\right) = E_v^{-1}\frac{\partial x}{\partial w}.
\end{aligned}
\tag{A.19}
$$

Note also that because of the property $q\frac{\partial \tilde{x}}{\partial q} = 0$, one obtains

$$
p\frac{\partial \tilde{x}}{\partial q} = (q-t)\frac{\partial \tilde{x}}{\partial q} = -t\frac{\partial \tilde{x}}{\partial q}. \tag{A.20}
$$

Using these two conditions in Equation (A.18) gives the expression in Equation (23).

We now turn to the derivation of the effective marginal tax rate. Note that the marginal rate of substitution between income and expenditure on commodity goods can be written as $s = \left(-\frac{b}{z}\right)_u = -E_z(q,u,z,w)$. This means that

$$
\begin{aligned}
\frac{\partial \left(E_w/E_v\right)}{\partial z} &= E_v^{-1}\left(E_{wz} - \frac{E_w}{E_v}E_{vz}\right) = E_v^{-1}\left(E_{wz} - \frac{E_w}{E_v}E_{vz}u_z\right) \\
&= E_v^{-1}s_w(q,z,u,w).
\end{aligned}
\tag{A.21}
$$

Also, because

$$p\frac{\partial \tilde{x}}{\partial z} = q\frac{\partial \tilde{x}}{\partial z} - (q-p)\frac{\partial \tilde{x}}{\partial z} = E_z - t(\frac{\partial x}{\partial z} + \frac{\partial x}{\partial b}E_z) = -\left(1 - t\frac{\partial x}{\partial b}\right)s - t\frac{\partial x}{\partial z}, \tag{A.22}$$

one can rewrite the first-order condition in Equation (24) as the rule in Equation (25).

A.2.5 Non-linear taxation with public provision

Under nonlinear taxation and public good provision the government maximizes its social welfare function subject to the incentive compatibility constraint $\left(v'(w) = -\frac{E_w}{E_v}\right)$ and the resource constraint $\left(\int(z - p\tilde{x} - \pi g)f\right)$. In addition to the presence of public provision g in the commodity demand, the cost of that provision πg needs to be taken into account in the resource constraint. The Lagrangian becomes:

$$\begin{aligned} L &= \int \left\{ O\left[e\left(\tilde{x}_c(z,v,w,g),g\right)\right] + \mu(z - p\tilde{x} - \pi g)f + \alpha v'(w) + \alpha\frac{E_w}{E_v} \right\} dw \\ &= \int \left\{ O\left[e\left(\tilde{x}_c(z,v,w,g),g\right)\right] + \mu(z - p\tilde{x} - \pi g)f - \alpha' v(w) + \alpha\frac{E_w}{E_v} \right\} dw \\ &\quad + \alpha(\infty)v(\infty) - \alpha(0)v(0). \end{aligned} \tag{A.23}$$

The first-order condition with respect to g is:

$$\int O'\left(\frac{\partial e}{\partial x_c}\frac{\partial \tilde{x}_c(z,v,w,g)}{\partial q} + \frac{\partial e}{\partial g}\right) f dw - \int \left(\mu\frac{\partial \tilde{x}(q,z,v,w)}{\partial g} f - \mu\pi f + \alpha\frac{\partial\left(E_w/E_v\right)}{\partial g}\right) dw = 0. \tag{A.24}$$

To derive the optimum condition for public provision of education services under non-linear taxation, note that

$$\begin{aligned} \frac{\partial\left(E_w/E_v\right)}{\partial g} &= E_v^{-1}\left(E_{wg} - E_w\frac{E_{vg}}{E_v}\right) = E_v^{-1}\left(E_{wg} + E_{vg}v'\right) \\ &= E_v^{-1} m_w(q,z,u,w), \end{aligned} \tag{A.25}$$

where $m = \frac{v_g}{v_b} = -E_g$. Further, since $t = 0$, we have:

$$p\frac{\partial \tilde{x}}{\partial g} = q\frac{\partial \tilde{x}}{\partial g} = E_g = -m. \tag{A.26}$$

Using these two formulae in Equation (A.23) leads to the rule in Equation (26).

A.3 Fair taxation

A.3.1 Income taxation

Individuals maximize their utility $u^i\left(e^i, x^i, l^i\right)$ subject to the budget constraint $(1-\tau)\, w^i l^i + b = x_a^i + x_c^i\left(e^i\right)$. The optimization leads to the indirect utility $v^i(1-\tau, b)$. The optimal policies maximize the money-metric measure of the worst-off person $\vartheta^o\left(\bar{w}, \bar{c}, v^i(1-\tau, b)\right)$. The government budget constraint remains intact.

The first-order conditions for optimal policies are:

$$\frac{\partial \vartheta^o}{\partial v^o}\frac{\partial v^o}{\partial(1-\tau)} + \mu\left(-\sum_i w^i l^i + \sum_i \tau w^i \frac{\partial l^i}{\partial(1-\tau)}\right) = 0 \tag{A.27}$$

$$\frac{\partial \vartheta^o}{\partial v^o}\frac{\partial v^o}{\partial b} + \mu\left(\sum_i \tau w^i \frac{\partial l^i}{\partial b} - N\right) = 0. \tag{A.28}$$

Dividing these equations side by side and utilizing Roy's identity gives:

$$w^o l^o = \frac{\sum_i w^i l^i - \sum_i \tau w^i \frac{\partial l^i}{\partial(1-\tau)}}{N - \sum_i \tau w^i \frac{\partial l^i}{\partial b}}. \tag{A.29}$$

Dividing the numerator and denominator of the right-hand side of Equation (A.29) by N and expressing labour supply effects in elasticity form leads to the equation:

$$\left(1 - \tau \sum_i \frac{w^i l^i}{Nb}\frac{\partial l^i}{\partial b}\frac{b}{l^i}\right) w^o l^o = \frac{\sum_i w^i l^i}{N} - \frac{\tau}{1-\tau}\sum_i \frac{w^i l^i}{N}\frac{\partial l^i}{\partial(1-\tau)}\frac{1-\tau}{l^i}, \tag{A.30}$$

which can be rewritten as:

$$A\theta^o = 1 - \frac{\tau}{1-\tau}\sum_i \theta^i \varepsilon^i_{1-\tau} \tag{A.31}$$

Here, $A \equiv 1 - \tau \sum_i \frac{w^i l^i}{Nb}\varepsilon^i_b > 1$, ε^i_x is the elasticity of labour supply with respect to x, and $\theta^i \equiv \frac{w^i l^i}{\sum_i w^i l^i}$. Equation (A.31) can be solved for the tax rate, given in Equation (31) in the main text.

A.3.2 Public provision

The optimal public provision condition in Equation (32) is directly analogous to the welfarist case. Equation (35) is achieved by redefining the net social marginal value of income to person i in the fairness model as $\rho^i_F = \frac{\beta^i_F}{\mu} + \tau w^i \frac{\partial l^i}{\partial b}$ and substituing into Equation (32).

A.3.3 Commodity taxation

Let the consumer price of good j be q_j with $q_j = 1 + t_j$. Ignoring public provision, the optimal policy maximizes $\vartheta^o = \vartheta^o(\bar{c}, \bar{w}, u^o(e^o, x^o, l^o))$, where the current choices by the individual maximize utility $u^o(e^o, x^o, l^o)$ subject to the budget constraint $w^o l^o + b \geq \sum_j q_j x_j^o + c^o(e^o)$. This gives the indirect utility $v^o(q, b)$ with q denoting the vector of consumer prices. The social welfare function now becomes $\vartheta^o = \vartheta^o(\bar{c}, \bar{w}, v^o(q, b))$. The government maximizes this with the budget constraint (as above in the welfarist case): $\sum_i \sum_j t_j x_j^i = Nb + R$, where the individual choices of consumption, labour supply, and education depend on tax rates t_j and the lump sum income b. The first-order conditions are:

$$\frac{\partial \vartheta^o}{\partial v^o} \frac{\partial v^o}{\partial q_k} + \mu \left(\sum_i x_k^i + \sum_i \sum_j t_j \frac{\partial x_j^i}{\partial q_k} \right) = 0 \tag{A.32}$$

$$\frac{\partial \vartheta^o}{\partial v^o} \frac{\partial v^o}{\partial b} + \mu \left(\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b} \right) - \mu N = 0. \tag{A.33}$$

Denote $\beta_F^o = \frac{\partial \vartheta^o}{\partial v^o} \frac{\partial v^o}{\partial b}$ for the worst-off person o. For all other $i \neq o$, $\beta_F^i = 0$. The direct marginal weight of a person in social welfare is 0 for all others than the worst-off person.[3] Utilising this, and the fact that $\frac{\partial v}{\partial q_k} = -\frac{\partial v}{\partial b} x_k$ for any indirect utility function, Equation (A.32) can be rewritten as:

$$-\sum_i \beta_F^i x_k^i + \mu \left(\sum_i x_k^i + \sum_i \sum_j t_j \frac{\partial x_j^i}{\partial q_k} \right) = 0. \tag{A.34}$$

Equation (A.33) can be rewritten as:

$$\sum_i \beta_F^i + \mu \left(\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b} \right) = \mu N. \tag{A.35}$$

Now Equations (A.34) and (A.35) are formally identical to their counterparts in the welfarist case. The tax rule in Equation (37) can be derived from Equation (A.35) using the Slutsky equation $\frac{\partial x_j^i}{\partial q_k} = \frac{\partial \tilde{x}_j^i}{\partial q_k} - x_k^i \frac{\partial x_j^i}{\partial b}$ and Slutsky symmetry.

In order to prove Equation (39), note first that the term in the right-hand side is obtained in a fashion similar to the commodity taxation case in the Appendix of Kanbur, Paukkeri, Pirttilä, and Tuomala (2018) by using $\frac{\sum_i \rho_F^i}{N} = 1$ (implied by Equation (A.35)).

[3] This is not true of the net social marginal utility of income of a person (see Equation (38)). The net social marginal utility of a person other than the worst-off is positive if the change in her consumption due to a higher lump sup transfer increases commodity tax revenue, and negative in the reverse case.

By definition,

$$N\mathrm{cov}\left(\rho_F^i, x_k^i\right) = \left(\frac{\beta^o}{\mu} + \sum_i t_j \frac{\partial x_j^o}{\partial b} - 1\right)\left(x_k^o - \bar{x}^k\right) \tag{A.36}$$
$$+ \left(\sum_i t_j \frac{\partial x_j^1}{\partial b} - 1\right)\left(x_k^1 - \bar{x}^k\right) + ... + \left(\sum_i t_j \frac{\partial x_j^{N-1}}{\partial b} - 1\right)\left(x_k^{N-1} - \bar{x}^k\right).$$

Use Equation (A.35) again to get:

$$\frac{\beta^o}{\mu} = N - \sum_j t_j \frac{\partial x_j^o}{\partial b} - ... \sum_j t_j \frac{\partial x_j^{N-1}}{\partial b}. \tag{A.37}$$

Substitute this in the covariance expression and note that $x_k^i - \bar{x}_k = x_k^i - x_k^o + x_k^o - \bar{x}_k$ to get the following expression:

$$N\mathrm{cov}\left(\rho_F^i, x_k^i\right) = \left(1 - \sum_i t_j \frac{\partial x^1}{\partial b}\right)\left(x_k^o - x_k^1\right) + ... + \left(1 - \sum_i t_j \frac{\partial x^{N-1}}{\partial b}\right)\left(x_k^o - x_k^{N-1}\right). \tag{A.38}$$

By rearranging this, one can separate out terms in x_k^o and the sum of other consumption levels. Requiring then the covariance to be negative results in Equation (39).