This paper is a draft submission to

Inequality—Measurement, trends, impacts, and policies

5–6 September 2014 Helsinki, Finland

Optimal Taxation and Public Good Provision for Poverty Minimization

Ravi Kanbur (Cornell University), Jukka Pirttilä (UNU-WIDER),
Matti Tuomala (University of Tampere) and Tuuli Ylinen (Aalto University)

First draft (August 2014)

Abstract

The purpose of the paper is to examine the determinants of optimal redistributive policies in the context of a developing country that can only implement linear tax policies due to administrative reasons. The optimal conditions for linear income taxation, commodity taxation and public provision of private and public goods are provided for the poverty minimization case, and the results are compared to those derived under a general welfarist objective function. These formulae capture the sufficient statistics that the governments need to pay attention to when designing poverty alleviation policies. The results also cover the cases where public provision of certain goods (such as education and health) serves to improve the capabilities of the citizens.

Contents

1 Introduction

Rising within-country inequality in many otherwise successful developing countries has become a key policy concern in global development debate. While some countries have very unequal inherent distributions (e.g. due to historical land ownership arrangements), in others the fruits of economic growth have been unequally shared. No matter what the underlying reason for the high inequality, often the only direct way for governments to affect the distribution of income is via redistributive tax and transfer systems. Clearly, public spending on social services has also an impact on the distribution of wellbeing, although some of the effects (such as skill-enhancing impacts from educational investment) only materialize over a longer time horizon.

Redistributive transfer systems have, indeed, proliferated in many developing countries, starting from Latin America and now spreading to low-income countries, including those in Sub-Saharan Africa.[1] In low-income countries, in particular, redistributive arrangements via transfers are still at an early stage, and they often consist of isolated programs. There is an urgent and well-recognized need to move away from scattered programs to more comprehensive tax-benefit systems.

This paper examines the optimal design of cash transfers, commodity taxes (or subsidies), the provision of public and private goods (such as education and housing), and financing them by a linear income tax. We build on the optimal income tax approach, which is extensively used in the developed country context[2], but much less applied for the design of redistributive systems in developing country circumstances. This approach, intitiated by Mirrlees (1971), allows for a rigorous treatment of efficiency concerns (e.g. the potentially harmful effect of distortionary taxation on employment) and redistributive objectives. Achieving the government's redistributive objectives is constrained by limited information: the social planner cannot directly observe individuals' income earning capacity, and therefore it needs to base its tax and transfer policies on observable variables, such as gross income. The most general formulation of optimal tax models apply non-linear tax schedules, but in a developing country context, using fully non-linear taxes is rarely feasible. In this paper we therefore limit the analysis to redistributive linear income taxes, which combine a lump-sum transfer with a proportional income tax, and which can be implemented by withholding at source if necessary.

In conventional optimal taxation models, the government's objective function is modelled as a social welfare function, which depends directly on individual utilities. We depart from this welfarist approach by presenting general non-welfarist tax rules, as in Kanbur, Pirttila, and Tuomala (2006), and, in particular, optimal tax and public good provision rules when the government is assumed to minimize poverty. We have chosen this approach as it resembles well the tone of much of policy discussion in developing countries, including the MDGs, where the

[1] For a recent treatment and survey, see Barrientos (2013).

[2] See IFS and Mirrlees (2011) for an influential application of optimal tax theory to policy analysis for rich countries.

objective is explicitly to reduce poverty rather than maximize wellbeing. Similarly, the discussion regarding cash transfer systems is often couched especially in terms of poverty alleviation. In all our analysis, we first present welfarist tax rules (which are mostly already available in the literature) to provide a benchmark to examine how applying poverty minimization as an objective changes the optimal tax and public service provision rules. We also deal with some extensions to existing models, which are motivated by the developing country context, such as the case where public provision affects the individuals' income-earning capacity, thus capturing (albeit at a very stylized way) possibilities to affect their capabilities.

Our paper is related to various strands of earlier literature. First, Kanbur, Keen, and Tuomala (1994) and Pirttila and Tuomala (2004) study optimal income tax and commodity tax rules, respectively, from the poverty alleviation point of view, but their papers build on the non-linear tax approach which is not well suited to developing countries. Second, the previous work on taxation and development, such as Gordon and Li (2009), Keen (2012) and Besley and Persson (2013), while clearly very relevant, has not concentrated much on the design of optimal redistributive systems. Finally, we follow the approach in Piketty and Saez (2013), who 'use the "sufficient statistics" approach whereby optimal tax formulas are derived and expressed in terms of estimable statistics including social marginal welfare weights capturing society's value for redistribution and labor supply elasticities capturing the efficiency costs of taxation' (Piketty and Saez (2013), p. 394). This sufficient statistic approach has proved very valuable for applied tax analysis, since it provides clear guidelines for the sort of empirical work that is needed to generate knowledge for implementing optimal tax rules. Piketty and Saez also emphasize how linear tax rules, while analytically more feasible, fit well with this approach as they lead to tax formulas that contain the same sufficient statistics than more complicated non-linear models. The linear tax rules, they argue, are robust to alternative specifications, and examining this forms part of our motivation: we study optimal linear tax policies, in our understanding for the first time, from the poverty minimization perspective.

The paper proceeds as follows. Section 2 examines optimal linear income taxation, while 3 turns to optimal provision rules for publicly provided private and public goods that are financed by such a linear income tax. Section 4 analyzes the combination of optimal linear income taxes and commodity taxation and asks under which conditions one should use differentiated commodity taxation if the government is interested in poverty minimization and also has optimal cash transfers at its disposal. Finally, conclusions are provided in Section 5.

2 Linear income taxation

2.1 Optimal linear income taxation under the welfaristic objective

In this section we give an overview of some of the models and results for optimal linear income taxation as they have been presented in the literature. We focus on the recent model by Piketty and Saez (2013), Tuomala (1985) modeling and the Dixit and Sandmo (1977) model.

The government collects a linear income tax τ, which it uses to finance a lump-sum transfer b, along with other exogenous public spending R. z^i denotes individual labour income ("wL"), and consumption equals $c^i = (1-\tau)z^i + b$, where superscript-i refers to individuals.

We start by following Piketty and Saez's modeling in defining a continuum of individuals, whose distribution is $\nu(i)$ (population size is normalized to one), and in using individual utility functions defined directly over consumption and leisure. The model also uses the government's budget constraint to define things in terms of the tax rate only. Individuals maximize their utility $u^i((1-\tau)z^i + b, z^i)$, and their FOC implicitly defines the Marshallian earnings function $z^i_u(1-\tau, b)$. Using this, aggregate earnings are $Z_u(1-\tau, b)$. The government's budget constraint $b + R = \tau Z_u(1-\tau, b)$ implicitly defines b as a function of τ, and consequently Z_u can also be defined solely as a function of τ: $Z(1-\tau) = Z_u(1-\tau, b(\tau))$. Z has elasticity $e = \frac{1-\tau}{Z}\frac{dZ}{d(1-\tau)}$.

To start, note that if the government only cared about maximizing tax revenue $\tau Z(1-\tau)$, it would set τ such that $\frac{\partial(\tau Z(1-\tau))}{\partial\tau} = 0$: $Z(1-\tau) - \tau\frac{dZ}{d(1-\tau)} = 0$. Using $\frac{\tau}{Z}\frac{dZ}{d(1-\tau)} = \frac{\tau}{1-\tau}e$, this gives

$$
\begin{aligned}
\frac{\tau^*}{1-\tau^*} &= \frac{1}{e} \\
\Leftrightarrow \tau^* &= \frac{1}{1+e}
\end{aligned}
\tag{2.1}
$$

Our reference point is a government concerned about social welfare. Taking into account that individual consumption is $c^i = (1-\tau)z^i + b = (1-\tau)z^i + \tau Z(1-\tau) - R$, its problem is to $max\, SWF = \int \omega^i W(u^i((1-\tau)z^i + \tau Z(1-\tau) - R, z^i)\, d\nu(i)$. Here ω is a Pareto weight and W is an increasing and concave transformation of utilities. The FOC $\frac{\partial SWF}{\partial\tau} = 0$ is:

$$\int \omega^i W_u \left[u^i_c\left(-z^i + (1-\tau)\frac{\partial z^i}{\partial\tau} + Z + \tau\frac{dZ}{d\tau}\right) + u^i_z\frac{\partial z^i}{\partial\tau}\right] \mathrm{d}\nu(i) = 0$$

which, using the individual's envelope condition, becomes:

$$\int \omega^i W_u u^i_c\left(-z^i + Z - \tau\frac{dZ}{d(1-\tau)}\right) \mathrm{d}\nu(i) = 0$$

Taking $Z - \tau\frac{dZ}{d(1-\tau)}$ out of the integrand and leaving it to the left-hand side we have on the right-hand side $\frac{\int \omega^i W_u u^i_c z^i\, \mathrm{d}\nu(i)}{\int \omega^j W_u u^j_c\, \mathrm{d}\nu(j)}$. Piketty and Saez define $\beta^i = \frac{\omega^i W_u u^i_c}{\int \omega^i W_u u^i_c d\nu(i)}$ as a normalized social marginal welfare weight for individual i, so that the term can be simplified to:

$$Z - \tau\frac{dZ}{d(1-\tau)} = \int \beta^i z^i\, \mathrm{d}\nu(i)$$

Using the definition of aggregate elasticity of earnings and defining $\bar{\beta} = \frac{\int \beta^i z^i d\nu(i)}{Z}$ as the average normalized social marginal welfare weight, weighted by labour incomes z^i (can also be interpreted as the ratio of the average income weighted by individual welfare weights β^i to the average income Z), we can rewrite this as:

$$1-\frac{\tau}{1-\tau}e \quad = \quad \bar{\beta}$$

According to Piketty and Saez, $\bar{\beta}$ "measures where social welfare weights are concentrated on average over the distribution of earnings". The social welfare maximizing tax rate is thus:

$$\begin{aligned} \frac{\tau^*}{1-\tau^*} &= \frac{1}{e}\left(1-\bar{\beta}\right) \\ \Leftrightarrow \tau^* &= \frac{1-\bar{\beta}}{1-\bar{\beta}+e} \end{aligned} \tag{2.2}$$

The welfare-maximizing tax rate is thus decreasing in both the average marginal welfare weight and the tax elasticity of aggregate earnings. A higher $\bar{\beta}$ reflects a lower taste for redistribution, and thus a lower desire to tax for redistributive reasons.

Piketty and Saez also note that (2.2) can be written in the form of $\tau^* = \frac{-cov\left(\beta^i, \frac{z^i}{Z}\right)}{-cov\left(\beta^i, \frac{z^i}{Z}\right)+e}$. If higher incomes are valued less (lower β) then the covariances are negative and the tax rate is positive. This is a similar formulation as in Dixit and Sandmo (1977), equation (20), where $\tau^* = -\frac{1}{\lambda}\frac{-cov\left(z^i,\mu^i\right)}{\frac{\partial \bar{z}}{\partial(1-\tau)}|_{comp.}}$ (here λ represents the government's budget constraint Lagrange multiplier and μ^i the individual's marginal utility of income, s.t. $U_c = \mu^i$). Here the numerator reflects the equity element and the denominator the efficiency component, similar as in (2.2).

Yet another expression for the optimal tax rule is available in Tuomala (1985). The government again has redistributive objectives represented by a Bergson-Samuelson functional $W\left(V^1, ..., V^N\right)$ with $W' > 0$, $W'' < 0$. Here we use indirect individual utility functions $V^i(1-\tau, b)$ instead of the direct utility functions u as previously, and denote the net-of tax rate $1-\tau = a$ to simplify the notation (e.g. subscript-a refers to the derivative with respect to the net-of-tax rate). We return to the case of a discrete distribution of N individuals. The government's problem is the same as before, to choose the tax rate τ and transfer b so as to maximize the social welfare function $\sum W\left(V^i(a,b)\right)$ under the budget constraint $(1-a)\sum z^i = Nb+R$. With λ denoting the multiplier associated with the budget constraint, the government's Lagrangian is $L = \sum W\left(V^i(a,b)\right) + \lambda\left((1-a)\sum z^i - Nb - R\right)$.

Using Roy's theorem, $V_a^i = V_b^i z^i$, the first order conditions with respect to a and b, respectively, are:

$$\sum \beta^i z^i \quad = \quad \lambda\left(\sum z^i - (1-a)\sum z_a^i\right) \tag{2.3}$$

$$\sum \beta^i \quad = \quad \lambda\left(N - (1-a)\sum z_b^i\right) \tag{2.4}$$

where $\beta^i = W_V V_b^i$ is the social marginal utility of income ($\Rightarrow W_V V_a^i = W_V V_b^i z^i = \beta^i z^i$).

Divide (2.3) by (2.4) to get:

$$\frac{\sum \beta^i z^i}{\sum \beta^i} = \frac{\sum z^i - (1-a)\sum z^i_a}{N - (1-a)\sum z^i_b} \tag{2.5}$$

Denote average income $\bar{z} = \frac{\sum z^i}{N}$ and welfare weighted average income $z(\beta) = \frac{\sum \beta^i z^i}{\sum \beta^i}$ to get:

$$z(\beta) = \frac{\bar{z} - (1-a)\bar{z}_a}{1 - (1-a)\bar{z}_b} \tag{2.6}$$

Multiply the government's revenue constraint by $\frac{1}{N}$ and define $g = \frac{R}{N}$ to get $(1-a)\bar{z} - b = g$, and totally differentiate:

$$\frac{db}{da}|_{gconst} = \frac{\bar{z} + (1-a)\bar{z}_a}{-\left[1 - (1-a)\bar{z}_b\right]} = -z(\beta) \tag{2.7}$$

That $z(\beta) = -\frac{db}{da}|_{gconst}$ tells us that welfare-weighted labour supply should be equal to the constant-revenue effect of tax rate changes in b.

By totally differentiating average labour income $\bar{z}$ and using (2.7), we have

$$\frac{d\bar{z}}{da}|_{gconst} = \bar{z}_a + \bar{z}_b\frac{db}{da}|_{gconst} = \bar{z}_a + \bar{z}_b z(\beta) \tag{2.8}$$

When we impose g as a constant we have to give up one of our degrees of freedom. Now the interpretation of $\frac{d\bar{z}}{da}|_{gconst}$ is then the effect on labour supply when a is changed, as is b, in order to keep tax revenue constant. Using (2.8) we can write (2.6):

$$z(\beta) - \bar{z} = (1-a)\frac{d\bar{z}}{da}|_{gconst} = -\tau\frac{d\bar{z}(1-\tau)\bar{z}}{d(1-\tau)(1-\tau)\bar{z}} \tag{2.9}$$

from which we get the optimal tax rate:

$$\frac{\tau^*}{1-\tau^*} = \frac{1}{e}\left(1 - \frac{z(\beta)}{\bar{z}}\right) \tag{2.10}$$

where $e = \frac{d\bar{z}}{d(1-\tau)}\frac{(1-\tau)}{\bar{z}}$ is the elasticity of total income with respect to the net-of-tax rate. Define $\Omega = \frac{z(\beta)}{\bar{z}}$, so that $I = 1 - \Omega$ is a normative measure of inequality or, equivalently, of the relative distortion arising from the second best tax system. Clearly Ω should vary between zero and unity. One would expect it to be a decreasing function of τ (given the per capita revenue requirement $g = R/N$). There is a minimum feasible level of τ for any given positive g, and of course g must not be too large, or no equilibrium is possible. Hence any solution must also satisfy $\tau > \tau_{min}$ if the tax system is to be progressive. That is, if the tax does not raise sufficient revenue to finance the non-transfer expenditure, R, the shortfall must be made up by imposing a poll tax ($b < 0$) on each individual. One would also expect the elasticity of labour supply with respect to the net-of-tax rate to be an increasing function of τ (it need not be).

We thus have the same result for optimal linear tax as in the model of Piketty and Saez.

We can rewrite (2.10) as $\tau^* = \frac{1-\Omega}{1-\Omega+e}$ (cf. equation (2.2)) to illustrate the basic properties of the optimal tax rate. Because $e \geq 0$ and $0 \leq \Omega < 1$, both the numerator and denominator are nonnegative. The optimal tax rate is thus between zero and one. The formula captures neatly the efficiency-equity trade off. τ decreases with e and Ω and we have the following general results: 1) In the extreme case where $\Omega = 1$, i.e. the government does not value redistribution at all, $\tau = 0$ is optimal. We can call this case libertarian. According to the libertarian view the level of disposable income is irrelevant (ruling out both basic income b, and other public expenditures, g, funded by the government). 2) If there is no inequality, then again $\Omega = 1$ and $\tau = 0$. There is no intervention by the government. The inherent inequality will be fully reflected in the disposable income. Furthermore, lumpsum taxation is optimal; $b = -g$ or $T = -b$. 3) We can call the case where $\Omega = 0$ as "Rawlsian" or maximin. The government maximizes tax revenue as optimal $\tau = \frac{1}{e}$, i.e. maximizes the basic income b (assuming the worst off individual has zero labour income). In fact, maximizing b can be regarded as a non-welfarist case, to which we now turn.

2.2 Optimal linear income taxation under non-welfarist objectives

A non-welfarist government is one that follows a different set of preferences than those employed by individuals themselves (this follows Kanbur, Pirttila, and Tuomala (2006)). Thus, instead of maximizing a function of individual utilities, the government has other, paternalistic objectives that go beyond utilities. A special case taken up in more detail below is the objective of minimizing poverty in the society. To be as general as possible, let us define a 'social evaluation function' (as in e.g. Kanbur, Pirttila, and Tuomala (2006)) as $S = \sum F(c^i, z^i)$, which the government maximizes instead of the social welfare function. $F(c^i, z^i)$ measures the social value of consumption c^i for a person with income z^i and can be related to $u(c^i, z^i)$ but is not restricted to it. Following Tuomala's model as above, given the instruments available, linear income tax τ, lump-sum grant b and other expenditure R the government maximizes the Lagrangian $L = \sum F(az^i + b, z^i) - \lambda\left((1-a)\sum z^i - Nb - R\right)$. The first-order conditions with respect to a and b are:

$$\begin{aligned}
\sum \left(F_c(z^i + az^i_a) + F_z z^i_a\right) + \lambda\left((1-a)\sum z^i_a - \sum z^i\right) &= 0 \\
\sum \left(F_c(1 + az^i_b) + F_z z^i_b\right) + \lambda\left((1-a)\sum z^i_b - N\right) &= 0
\end{aligned}$$

Dividing the first equation with the second and dividing through the right hand side with N we get:

$$\frac{\sum \left(F_c(z^i + az^i_a) + F_z z^i_a\right)}{\sum \left(F_c(1 + az^i_b) + F_z z^i_b\right)} = \frac{\bar{z} - (1-a)\bar{z}_a}{1 - (1-a)\bar{z}_b}$$

Define $\frac{\sum\left(F_c(z^i+az_a^i)+F_z z_a^i\right)}{\sum\left(F_c(1+az_b^i)+F_z z_b^i\right)} \equiv \tilde{F}$, which reflects the relative impact of taxes and transfers on the social evaluation function. Using this, and following the same steps as in the previous section, the optimal tax rate becomes:

$$\frac{\tau^*}{1-\tau^*} = \frac{1}{e}\left(1 - \frac{\tilde{F}}{\bar{z}}\right) \tag{2.11}$$

The result resembles (2.10). In addition to efficiency considerations via the term $\frac{1}{e}$, both entail a term that measures the relative benefits of taxes and transfers, in the welfarist case via welfare-weighted income $z(\beta)$, in the non-welfarist case via $\tilde{F}$, the relative impacts on the social evaluation function. However, the non-welfarist optimal tax rate might differ from the welfarist rate due to the F_z terms in $\tilde{F}$. The signs and magnitudes of F_c and F_z and thus $\tilde{F}$ depend on the specific objective of the government, i.e. the shape of F. Let us consider the specific case of poverty minimization below.

2.2.1 Special case: Poverty minimization

Now let us derive the optimal linear tax results for a government whose objective is to minimize poverty in the society. The instruments available to the government are the same, τ and b, and other exogenous expenditure is R. Note first that the revenue-maximizing tax rate in (2.1) is in fact equivalent to the tax rate obtained from a maxi-min objective function, since when the government only cares about the poverty (consumption) of the poorest individual, its only goal is to maximize redistribution to this individual, i.e. maximize tax revenue.

Let us first define the objective function of the government explicitly. Poverty is defined as deprivation of individual consumption c^i relative to some desired level $\bar{c}$ and measured with a deprivation index $D\left(c^i, \bar{c}\right)$, such that $D > 0 \, \forall \, c \in [0, \bar{c})$ and $D = 0$ otherwise, and $D_c < 0$, $D_{cc} > 0 \, \forall \, c \in [0, \bar{c})$, as in Pirttila and Tuomala (2004). Continuing with the model of Tuomala, the social evaluation function $F(c^i, z^i)$ becomes $D\left(c^i, \bar{c}\right)$ and the objective function thus becomes $min \; P = \sum D\left(c^i, \bar{c}\right)$. Now $F_c = D_c$ and $F_z = 0$, so $\tilde{F} = \tilde{D} = \frac{\sum D_c\left(z^i+az_a^i\right)}{\sum D_c\left(1+az_b^i\right)}$ and the optimal tax rule becomes:

$$\frac{\tau^*}{1-\tau^*} = \frac{1}{e}\left(1 - \frac{\tilde{D}}{\bar{z}}\right) \tag{2.12}$$

Since now $F_z = 0$, the result is closer to (2.10) than (2.11) was, and is likely to move in the same direction. Here $\tilde{D}$ describes the relative efficiency of taxes and transfers in reducing deprivation. Both the numerator and denominator of $\tilde{D}$ depend on D_c, so the difference in the the relative efficiency of the two depends on z_a^i and z_b^i. The more people react to taxes (relative to transfers) by earning less, the higher is $\tilde{D}$ and the lower should the tax rate be. In (2.10), the higher is the social value of income, the higher is $z(\beta)$ and the lower should the tax rate be.

We can also rewrite $\tilde{D}$, using $a = 1-\tau$, as: $\frac{\sum D_c\left(z^i+(1-\tau)\frac{\partial z^i}{\partial(1-\tau)}\right)}{\sum D_c\left(1+(1-\tau)z^i_b\right)} = \frac{\sum D_c\left(1+\frac{(1-\tau)}{z^i}\frac{\partial z^i}{\partial(1-\tau)}\right)z^i}{\sum D_c\left(1+(1-\tau)z^i_b\right)} = \frac{\sum D_c(1+e)z^i}{\sum D_c\left(1+(1-\tau)z^i_b\right)}$. Thus the $\tilde{D}$ in the optimal tax result (2.12) entails a further efficiency consideration, lowering optimal tax rates to induce the poor to work more. Kanbur, Keen, and Tuomala (1994) find a similar result in their nonlinear poverty-minimizing tax model. Here, however, we are restricted to lower the tax on everyone instead of on only the poorest individuals.

Let us now consider the poverty-minimizing results in the Piketty-Saez model. Given the government's instruments, consumption is $c^i = (1-\tau)z^i + b = (1-\tau)z^i + \tau Z(1-\tau) - R$. The poverty-minimization objective in the continuous case thus reads:

$$\begin{aligned} min\, P &= \int D\left(c^i, \bar{c}\right)\, \mathrm{d}\nu(i) \\ &= \int D\left((1-\tau)z^i + \tau Z(1-\tau) - R, \bar{c}\right)\, \mathrm{d}\nu(i) \end{aligned} \tag{2.13}$$

The optimal tax rate is found from the government's FOC, $\frac{\partial P}{\partial \tau} = 0$:

$$\begin{aligned} &\int D_c\left(-z^i + (1-\tau)\frac{\partial z^i}{\partial \tau} + Z + \tau\frac{dZ}{d\tau}\right)\, \mathrm{d}\nu(i) = 0 \\ \Leftrightarrow &\int D_c\left(-z^i - (1-\tau)\frac{\partial z^i}{\partial(1-\tau)} + Z - \tau\frac{dZ}{d(1-\tau)}\right)\, \mathrm{d}\nu(i) = 0 \end{aligned} \tag{2.14}$$

We will follow Piketty and Saez's model and define a "normalized marginal deprivation weight", similar to the normalized social marginal welfare weight β^i in the original model. Let $\beta^i = \frac{D_c}{\int D_c d\nu(j)}$ be this deprivation weight. We can interpret β^i as the effect on the deprivation measure of increasing disposable income marginally from level c^i, relative to the sum of all deprivation changes over the population (remember that $D_c < 0$ for $c < \bar{c}$ and $D_c = 0$ otherwise – β^i is thus positive for all $c < \bar{c}$ and decreasing in c). Using this definition, $\left(Z - \tau\frac{dZ}{d(1-\tau)}\right)\int D_c\, \mathrm{d}\nu(i) = \int D_c\left(z^i + (1-\tau)\frac{\partial z^i}{\partial(1-\tau)}\right)\, \mathrm{d}\nu(i)$ can be written as:

$$Z - \tau\frac{dZ}{d(1-\tau)} = \int \beta^i\left(z^i + (1-\tau)\frac{\partial z^i}{\partial(1-\tau)}\right)\, \mathrm{d}\nu(i) \tag{2.15}$$

Using the definition of the elasticity of individual labour earnings $e^i_c = \frac{1-\tau}{z^i}\frac{\partial z^i}{\partial(1-\tau)}$, we have $(1-\tau)\frac{\partial z^i}{\partial(1-\tau)} = z^i e^i_c$ and using elasticity of aggregate earnings $e = \frac{1-\tau}{Z}\frac{dZ}{d(1-\tau)}$ we have $Z - \tau\frac{dZ}{d(1-\tau)} = 1 - \frac{\tau}{1-\tau}e$ and we can rewrite the above as:

$$Z\left(1 - \frac{\tau}{1-\tau}e\right) = \int \beta^i\left(z^i + z^i e^i_c\right)\, \mathrm{d}\nu(i) \tag{2.16}$$

Define, analogous to Piketty-Saez, $\bar{\beta} = \frac{\int \beta^i z^i d\nu(i)}{Z} \left(= \frac{\int D_c z^i d\nu(i)}{Z \int D_c d\nu(j)} \right)$, an average normalized deprivation weight, weighted by labour incomes (can also be interpreted as average labour income weighted by individual deprivation weights). In addition, to simplify the notation, define $\bar{\beta}^e = \frac{\int \beta^i z^i e_c^i d\nu(i)}{Z} \left(= \frac{\int D_c z^i e_c^i d\nu(i)}{Z \int D_c d\nu(j)} \right)$, which describes average labour incomes weighted by their corresponding individual elasticities and deprivation weights. This can be interpreted as a combined deprivation and efficiency effect. Using these definitions, we get an optimal tax rule that resembles the welfaristic one in (2.2):

$$\begin{aligned} \frac{\tau^*}{1-\tau^*} &= \frac{1}{e}\left(1 - \bar{\beta} - \bar{\beta}^e\right) \\ \Leftrightarrow \tau^* &= \frac{1 - \bar{\beta} - \bar{\beta}^e}{1 - \bar{\beta} - \bar{\beta}^e + e} \end{aligned} \tag{2.17}$$

As in the welfaristic setting, the more elastic average earnings are to taxation, the lower is the optimal tax rate (a regular efficiency effect). The optimal poverty-minimizing tax rate is decreasing in the average deprivation weight $\bar{\beta}$, as a higher taste for redistribution towards the materially deprived implies a lower $\bar{\beta}$ and thus higher taxation for redistributive purposes. The effect is analogous to the welfaristic tax rate, of course with slightly different definitions for $\bar{\beta}$.

The new term $\bar{\beta}^e$ can be interpreted as a combined deprivation weight and efficiency effect. The elasticity term implicit in $\bar{\beta}^e$ takes into account the incentive effects of taxation on working and works to reduce τ^*. To avoid discouracing the poor from working, their tax rates should be lower. But because the tax instrument is forced to be linear, tax rates are then lowered for everyone, as we found in the Tuomala model in equation (2.12). The value of $\bar{\beta}^e$ depends on the relationship of the individual earnings elasticities and income: if the elasticity is the same across income levels, there is just a level effect moving from $\bar{\beta}$ to $\bar{\beta}^e$; however if the elasticity were higher for more deprived individuals, for example, $\bar{\beta}^e$ would most likely be higher than under a flat elasticity. This would indicate towards a lower tax rate in order to avoid discouraging the poorest from working. However, whether $\bar{\beta}^e$ is high or low does not depend only on the shape of the elasticity but also on the shape of the deprivation weights, which also affect $\bar{\beta}$.

Note that in the general non-welfaristic case we wouldn't get as neat a result for the optimal tax rate due to the term F_z, which in the general case does not disappear as here. This means that we would have a term $-\frac{F_z}{Z \int F_c \, \mathrm{d}\nu(i)}$ in the numerator and denominator of equation (2.17).

Finally, the third way for expressing the optimal tax rule in the case of poverty minimization is one following the Dixit and Sandmo (1977) formulation. It can be shown that the poverty minimizing tax rate can also be written as

$$\tau^* = -\frac{1}{\lambda} \frac{cov\left(D_c, z^i\right) + \frac{1}{N}\sum D_c a \tilde{z}_a^i + cov\left(D_c a z_b^i, z^i\right)}{\frac{1}{N}\sum \tilde{z}_a^i}. \tag{2.18}$$

In this expression, the denominator is the same as in equation (20) of Dixit and Sandmo (1977) presented in Section 2.2, that is, the average derivative of compensated labour supply with respect to the net-of-tax rate. In the numerator, the first term measures the strenght of the association between income and poverty impact: when the association between overall poverty and small income is strong (this would the case of squared poverty gap), the tax should be high so that it will finance a sizable lump-sum transfer. If the association is weaker (as in the headcount rate), the tax rate is optimally smaller. The second and the third terms in the numerator are new. They measure the indirect effects from changes in the tax rate on labor supply. Here $\tilde{z}$ is the compensated (Hicksian) labour supply. The greater is the reduction in the labor supply following an increase in the tax rate (it is the compensated change as the tax increase is linked with a simultaneous increase in the lump-sum transfer), the smaller should the tax rate be in order to avoid increases in deprivation arising from lower earned income. The last two terms in the numerator are closely linked with a formulation $D_c(1-\tau)\frac{\partial \bar{z}}{\partial q}|_{comp}$, where the idea is that the the last covariance term serves as a corrective device for the mean impact of taxes on labor supply (similarly as in the denominator in the original Dixit-Sandmo formulation).

To summarize, the nonwelfarist tax rules differs from the welfarist one, depending on the definition of nonwelfarism in question (the F_c and F_z terms). However, when we take poverty minimization as the specific case of nonwelfarism, the tax rules are quite similar to welfarist ones. The basic difference is that equity is not considered in welfare terms but in terms of poverty reduction effectiveness. A more notable difference arises from efficiency considerations. With linear taxation, taking into account labour supply responses means that everybody's tax rate is affected, instead of just the target group's. If we want want to induce the poor to work more to reduce their poverty, we need to lower everyone's tax rate. The welfarist linear tax rule does not take this into account. It is not however possible to state that under poverty-minimization tax rates are optimally lower than under welfare maximization, since we cannot directly compare the welfare and deprivation terms. However there is an additional efficiency consideration involved under poverty minimization. Nonlinear tax rules of course make it possible to target lower tax rates on the poorer individuals, but in a developing country context with lower administrative capacity this is not necessarily possible, and such considerations affect everyone's tax rate.

3 Public good provision with linear income taxes

3.1 Optimal public provision under the welfaristic objective

Let us first extend the welfaristic Piketty-Saez model of linear taxation to entail provision of public goods. The government offers a universal pure public good G, which enters individual utilities separately from other consumption x. The government's goal function is then: $SWF = \int \omega^i W\left(u^i\left(\left(1-\tau\right)z^i + \tau Z(1-\tau) - R - pG, G, z^i\right)\right) \mathrm{d}\nu(i)$, where p is the price of the public good. The FOC for τ is as before, and the FOC for public good provision G is $\int \omega^i W_u \left(u_G^i + u_x^i \left((1-\tau)\frac{\partial z^i}{\partial G} + \tau \frac{dZ}{dG} - p\right)\right) \mathrm{d}\nu(i) = 0$, which gives us:

$$\frac{\int \omega^i W_u \left(u_G^i + u_x^i \left((1-\tau)\frac{\partial z^i}{\partial G}\right)\right) \mathrm{d}\nu(i)}{\int \omega^i W_u u_x^i \,\mathrm{d}\nu(i)} \quad = \quad p - \tau \frac{dZ}{dG} \tag{3.1}$$

The left-hand side relates the welfare gains of public good provision (a direct (u_G) and indirect effect ($u_x \left((1-\tau)\frac{\partial z^i}{\partial G}\right)$ via labour supply reactions)) to the welfare gains of directly increasing consumption (cash transfers) and the right-hand side relates the costs of providing the public good (both its price and the effect it has on tax revenue) to the costs of directly increasing consumption (equal to 1 in this model).[3]

To include public good provision to Tuomala's model, the government's Lagrangian becomes $L = \sum W\left(V^i(a,b,G)\right) + \lambda\left((1-a)\sum z^i - Nb - NpG - R\right)$, where p is the producer price of the public good. The producer price of private consumption is normalised to 1. Let us now define the marginal willingness to pay for the public good by the expression $\sigma = \frac{V_G}{V_b}$. Maximizing the Lagrangean, the first order conditions with respect to b and G give:

$$\sum \beta^i \quad = \quad \lambda\left(N - (1-a)\sum z_b^i\right) \tag{3.2}$$

$$\sum W_V V_G^i \quad = \quad \lambda\left(Np - (1-a)\sum z_G^i\right) \tag{3.3}$$

Dividing (3.3) by (3.2) we obtain

$$\frac{\sum \beta^i \sigma^i}{\sum \beta^i} \quad = \quad \frac{p - (1-a)\bar{z}_G}{1 - (1-a)\bar{z}_b} \tag{3.4}$$

where we can define $\sigma^* = \frac{\sum \beta^i \sigma^i}{\sum \beta^i}$ as the welfare weighted average marginal rate of substitution between public good and income for individual i.This public good provision rule is a different version of a modified Samuelson rule. It equates the welfare weighted sum of MRS to the relative cost of providing the public good compared to income transfers, and takes into account the revenue impacts of both. The term $\tau\bar{z}_G$ describes the revenue effects of public good provision, which can be in practice very important. A revenue gain arising from provision of the public good strengthens the case for it. The equation can be further rewritten as

$$p = \sigma^* - \tau\left(\sigma^* \bar{z}_b - \bar{z}_G\right) \tag{3.5}$$

If, for example, labour supply is independent of the level of public good provision, $\bar{z}_G = 0$

[3] In equation (3.1), we could define a normalized marginal social welfare weight, similar as before, $\beta^i = \frac{\omega^i W_u u_x^i}{\int \omega^i W_u u_x^i d\nu(i)}$ to get $\frac{\int \omega^i W_u u_G^i \,\mathrm{d}\nu(i)}{\int \omega^i W_u u_x^i \,\mathrm{d}\nu(i)} + \int \beta^i (1-\tau)\frac{\partial z^i}{\partial G}\,\mathrm{d}\nu(i) = p + \tau\frac{dZ}{dG}$.

and $\bar{z}_b < 0$, and if σ^* is positive, then the second term in (3.5) would make p higher than the welfare weighted aggregate marginal rate of substitution.

3.2 Optimal provision of public goods under poverty minimization

Now consider a non-welfarist government interested in minimizing poverty. The public good G which it offers, enters the deprivation index separately from other, private consumption x: $D\left(x, G, \bar{x}, \bar{G}\right)$. The government still offers a lump-sum cash transfer b as well, and finances its expenses with the linear income tax τ. Using our version of the Piketty-Saez model, individual private consumption is then $x = (1-\tau)z^i + b = (1-\tau)z^i + \tau Z(1-\tau) - R - pG$, where p is the price of the public good. The government's problem is then:

$$\begin{aligned} min\ P &= \int D\left(x^i, G, \bar{x}, \bar{G}\right)\, \mathrm{d}\nu(i) \\ &= \int D\left((1-\tau)z^i + \tau Z(1-\tau) - R - pG, G, \bar{x}, \bar{G}\right)\, \mathrm{d}\nu(i) \end{aligned} \tag{3.6}$$

The first-order condition for optimal tax τ is unchanged, and the FOC for public good provision is $\int \left[D_G + D_x\left((1-\tau)\frac{\partial z^i}{\partial G} + \tau\frac{dZ}{dG} - p\right)\right] \mathrm{d}\nu(i) = 0$, which gives us the public provision rule:

$$\frac{\int \left(D_G + D_x(1-\tau)\frac{\partial z^i}{\partial G}\right) \mathrm{d}\nu(i)}{\int D_x\, \mathrm{d}\nu(i)} = p - \tau\frac{dZ}{dG} \tag{3.7}$$

In the numerator of the left-hand side, the first term is the direct deprivation effect of G and the second term captures the indirect deprivation effect, operating via the labour supply impacts of the public good, which affect the level of private consumption x. These impacts are scaled by the poverty alleviation impact of private consumption itself (the impact of a cash transfer). The right hand side reflects the costs of public good provision: besides the direct cost of the good there is an indirect tax revenue effect operating through labour supply. The condition is directly comparable to (3.1) because even though the welfaristic case relies on utilities, in the FOC for G no envelope condition is evoked (results (2.2) and (2.17) differ from each other because the envelope condition is evoked in the former and not in the latter). The only difference between equations (3.1) and (3.7) is that the utility and welfare weight terms are exchanged for deprivation terms.

Now consider the provision of a quasi-private good, such that in addition to the publicly provided amount, individuals can purchase ("top-up") the good themselves as well. The good is denoted s and its total amount consists of private purchases h and public provision G: $s = G + h$. In addition to good s, individuals consume other private goods denoted x. The individual budget constraint is thus $c^i = x^i + ph^i = (1-\tau)z^i + \tau Z(1-\tau) - R - pG$. Deprivation is determined in terms of consumption of x and s, so the objective is:

$$
\begin{aligned}
min\ P &= \int D\left(x^i, s^i, \bar{x}, \bar{s}\right) \mathrm{d}\nu(i) \\
&= \int D\left((1-\tau)z^i + \tau Z(1-\tau) - R - pG - ph^i, s^i, \bar{x}, \bar{s}\right) \mathrm{d}\nu(i) \qquad (3.8)
\end{aligned}
$$

The FOC for public good provision G is $\int \left[D_x \left((1-\tau)\frac{\partial z^i}{\partial s}\frac{\partial s}{\partial G} + \tau\frac{dZ}{dG} - p - p\frac{\partial h^i}{\partial G}\right) + D_s\frac{\partial s^i}{\partial G}\right] \mathrm{d}\nu(i) = 0$, which gives the public provision rule:

$$
\frac{\int \left[D_x \left((1-\tau)\frac{\partial z^i}{\partial s}\frac{\partial s}{\partial G} - p\frac{\partial h^i}{\partial G}\right) + D_s\frac{\partial s^i}{\partial G}\right] \mathrm{d}\nu(i)}{\int D_x \,\mathrm{d}\nu(i)} = p - \tau\frac{dZ}{dG} \qquad (3.9)
$$

The result is analogous to the pure public good result in (3.7), with the difference that now the impact G has on poverty depends on whether public provision fully crowds out private purchases of the good (i.e. $\frac{dh}{dG} = -1 \Leftrightarrow \frac{ds}{dG} = 0$) or not (i.e. $\frac{dh}{dG} = 0 \Leftrightarrow \frac{ds}{dG} = 1$). If there is full crowding out, public provision of G has no impact on the consumption of s and consequently no impact on poverty. If there is no crowding out however, the FOC becomes

$$
\frac{\int \left[D_x \left((1-\tau)\frac{\partial z^i}{\partial s}\right) + D_s\right] \mathrm{d}\nu(i)}{\int D_x \,\mathrm{d}\nu(i)} = p - \tau\frac{dZ}{dG} \qquad (3.10)
$$

which is the same as in the case of a pure public good in equation (3.7).

Using Tuomala's model, using the deprivation index as before, defined over consumption of the public good G and other private consumption x, we can divide the government's first order condition for G with that of b to get the following relationship:

$$
D^* = \frac{p - (1-a)\bar{z}_G}{1 - (1-a)\bar{z}_b} \qquad (3.11)
$$

where $D^* = \frac{\sum D_G + \sum D_x a z_G^i}{\sum D_x \left(1 + a z_b^i\right)}$ captures the efficiency of the public good in reducing deprivation relative to the income transfer (because D_G, $D_x < 0$, $D^* > 0$). This rule already differs considerably from standard modified Samuelson rules, reflecting instead of MRS the direct poverty reduction impact of the public good and its indirect impact via labour supply and thus consumption effects. The right hand side shows the relative cost, just as in (3.4). The tax revenue impacts are important to determining the optimal level of public good provision and income transfers. Note also the indirect effect on poverty reduction in D^* via labour supply effects. We can further rewrite this to achieve comparability with equation (3.5):

$$
p = D^* - \tau\left(D^* \bar{z}_b - \bar{z}_G\right) \qquad (3.12)
$$

This shows how the price of the public good relates to the relative deprivation efficiency of public good provision. Using the same example as in the context of (3.5), if $\bar{z}_G = 0$ and

$\bar{z}_b < 0$, the price p of the public good would be higher than its relative efficiency in eliminating deprivation.

Suppose now that the consumers' welfare does not directly depend on the public good provision but the public good can have a productivity increasing impact. An example could be publicly provided education services that affect individuals' productivity via the wage rate. We therefore suppose that the direct impact of the public good on deprivation cancels out (i.e. $D_G = 0$), whereas the wage rate becomes an increasing function of G, i.e. $w'(G) > 0$ (denoting $z = w(G)L$). This means that the expression in (3.11) is written as

$$\frac{\sum D_x a \left(w\frac{\partial L}{\partial G} + w'L\right)}{\sum D_x \left(1 + aw\frac{\partial L}{\partial b}\right)} = \frac{p - (1-a)\sum \left(w\frac{\partial L}{\partial G} + w'L\right)}{1 - (1-a)\sum w\frac{\partial L}{\partial b}}. \tag{3.13}$$

According to this rule, even if labor supply would not react to changes in the public good provision, public good provision would still be potentially desirable through its impact on the wage rate. In this way, public good provision can be interpreted as increasing the capability of the individuals to earn a living wage, which serves as poverty reducing tool, and which can in some cases be a more effective way to reduce poverty rather than direct cash transfers. The optimality depends on the relative strength of $w'(G) > 0$ versus the direct impact of the transfers.

To summarize, the welfaristic public provision rule, when public goods are financed with linear income taxes and supplemented with lump-sum transfers, differs from the standard modified Samuelson rule. It equates a welfare-weighted sum of MRS to the marginal cost where tax revenue impacts are taken into account. Indirect effects of public provision (through labour supply decisions and thus private consumption) are incorporated. The poverty-minimizing public provision rule however replaces the welfare-weighted sum of MRS with the relative marginal returns to deprivation reduction. Here the "MRS" term measures how well public good is translated to reduced poverty (incorporating indirect effects as well), relative to private consumption. Finally, when the public good has positive effects on productivity, its provision can be desirable even if it would not have any direct impact on poverty.

4 Commodity taxation with linear income taxes

4.1 Optimal commodity taxation with linear income tax under the welfaristic objective

This section considers the possibility that the government also uses commodity taxation (subsidies) to influence consumers' welfare. We follow the modeling of Diamond (1975). Unlike the analysis above, there are J consumer goods x_j instead of just two. Working with many goods is used to be able to more clearly describe the condtions under which uniform commodity taxation occurs at the optimum. The governments levies a tax t_j on the consumption of good x_j, so that its consumer price is $q_j = p_j + t_j$, where p_j represents the producer price (a commodity subsidy would be reflected by $t_j < 0$). Let q denote the vector of all consumer prices. In addition, the government can use a lump-sum transfer, b. Note that in this exposition, leisure is the untaxed numeraire commodity. Alternatively, one could also imply a linear tax on labor supply as above and treat one of the consumption goods as untaxed numeraire. However, choosing leisure as the numeraire makes the exposition easier. Thus, the consumer's budget constraint is $\sum_j q_j x_j = z + b$. The Lagrangean of the government's optimization problem is the following:

$$L = \sum_i W\left(V^i(b,q)\right) + \lambda\left(\sum_i \sum_j t_j x_j^i - Nb - R\right) \tag{4.1}$$

The first-order conditions with respect to b and q_k are

$$\sum_i \beta^i \quad + \lambda \textstyle\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b} - \lambda N \quad = 0 \tag{4.2}$$

$$-\sum_i \beta^i x_k^i \quad + \lambda \textstyle\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial q_k} + \lambda \sum_i x_k^i \quad = 0 \tag{4.3}$$

where in (4.3) use has been made of Roy's identity, i.e. $\frac{\partial V^i}{\partial q_k} = -\frac{\partial V^i}{\partial b} x_k^i$. Now define, following Diamond (1975)

$$\gamma^i = \beta^i + \lambda \sum_j t_j \frac{\partial x_j^i}{\partial b} \tag{4.4}$$

as the net social marginal utility of income for person i. This notion takes into account the direct marginal social gain, β^i , and the tax revenue impact arising from commodity demand changes. This means that (4.2) can we rewritten as

$$\frac{\sum_i \gamma^i}{N} = \lambda, \tag{4.5}$$

implying that the average net social marginal utility of income must equal the shadow price of budget revenues at the optimum. Next use the definition of γ and the Slutsky equation for the commodity demand

$$\frac{\partial x_j^i}{\partial q_k} = \frac{\partial \tilde{x}_j^i}{\partial q_k} - x_k^i \frac{\partial x_j^i}{\partial b}$$

where $\tilde{x}_j^i$ denotes the compensated (Hicksian) demand for good x_j^i, in (4.3) to get

$$\sum_i \sum_j t_j \frac{\partial \tilde{x}_j^i}{\partial q_k} = \frac{1}{\lambda} \sum \left(\gamma^i - \lambda\right) x_k^i. \tag{4.6}$$

The covariance between γ^i and the demand of the good x_k can be written as (using (4.5))

$$cov(\gamma^i, x_k^i) \;=\; \frac{\sum_i \gamma^i x_k^i}{N} - \frac{\sum_i \gamma^i}{N}\frac{\sum_i x_k^i}{N} = \frac{\sum_i \gamma^i x_k^i}{N} - \lambda\frac{\sum_i x_k^i}{N}.$$

Using Slutsky symmetry, equation (4.3) can therefore be written as a covariance rule

$$\frac{1}{N}\sum_i\sum_j t_j \frac{\partial \tilde{x}_k^i}{\partial q_j} = \frac{1}{\lambda}cov(\gamma^i, x_k^i). \tag{4.7}$$

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 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.

The key policy question is whether or when uniform commodity taxes are optimal, or, in other words, when would a linear income tax combined with an optimal demogrant be sufficient to reach the society's distributional goals at the smallest cost. Deaton (1979) shows how weakly separable consumption and leisure and linear Engel curves are sufficient conditions for the optimality of uniform commodity taxes. These requirements are quite stringent and unlikely to hold in practice; however, the economic importance / magnitude they imply is unclear. If implementing differentiated commodity taxation entails significant administrative costs, they may easily outweigh the potential benefits of distributional goals, and that is why economists have typically been quite sceptical about non-uniform commodity taxation when applied to practical tax policy.

4.2 Optimal commodity taxation with linear income tax under poverty minimization

Poverty could be measured in many ways with multiple commodity goods: the government may care about overall consumption, the consumption of some of the goods (those that are in the basket used to measure poverty) or then it cares about both the overall consumption and the relative share of different kinds of consumption goods (such as merit goods). Here we examine the simplest set-up where deprivation only depends on the disposable income, $c = z + b$. Using the consumer's budget constraint, this is equal to the overall consumption level, $\sum q_j x_j^i$.

The government thus minimizes the sum of the poverty index $D\left(\sum q_j x_j^i, \bar{c}\right)$, and the first-order conditions with respect to b and q_k are:

$$\sum_i D_c \sum_j q_j \frac{\partial x_j^i}{\partial b} \quad +\lambda \textstyle\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b} - \lambda N \quad = 0 \tag{4.8}$$

$$\sum_i D_c x_k^i + \sum_i D_c \sum_j q_j \frac{\partial x_j^i}{\partial q_k} \quad +\lambda \textstyle\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial q_k} + \lambda \sum_i x_k^i \quad = 0 \tag{4.9}$$

Using the Slutsky equation in equation 4.9 and dividing by N leads to

$$\begin{aligned}&\tfrac{1}{N} \textstyle\sum_i D_c x_k^i + \tfrac{1}{N} \sum_i D_c \sum_j q_j \left(\frac{\partial \tilde{x}_j^i}{\partial q_k} - x_k^i \frac{\partial x_j^i}{\partial b} \right) \\ &+ \tfrac{\lambda}{N} \textstyle\sum_i \sum_j t_j \left(\frac{\partial \tilde{x}_j^i}{\partial q_k} - x_k^i \frac{\partial x_j^i}{\partial b} \right) + \tfrac{\lambda}{N} \sum_i x_k^i = 0 \end{aligned} \tag{4.10}$$

Multiplying equation (4.8) by $\sum_i x_k^i / N^2$ and adding it with equation (4.10) gives

$$\begin{aligned}&\tfrac{1}{N} \textstyle\sum_i D_c x_k + \tfrac{1}{N} \sum_i D_c \sum_j q_j \frac{\partial \tilde{x}_j^i}{\partial q_k} - \tfrac{1}{N} \sum_i \sum_j D_c q_j x_k^i \frac{\partial x_j^i}{\partial b} \\ &+ \tfrac{1}{N} \tfrac{\sum_i D_c}{N} \textstyle\sum_j q_j \frac{\partial x_j^i}{\partial b} \sum_i x_k^i + \tfrac{\lambda}{N} \sum_i \sum_j t_j \frac{\partial \tilde{x}_j^i}{\partial q_k} \\ &- \tfrac{\lambda}{N} \textstyle\sum_i \sum_j t_j x_k^i \frac{\partial x_j^i}{\partial b} + \tfrac{1}{N} \tfrac{\lambda}{N} \sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b} \sum_i x_k^i \qquad = 0 \end{aligned} \tag{4.11}$$

Rearranging, using the definition of covariance and Slutsky symmetry allows us to write this expression as

$$\begin{aligned}\frac{1}{N} \sum_i \sum_j t_j \frac{\partial \tilde{x}_k^i}{\partial q_j} \quad = \quad & -\frac{1}{\lambda} \left[\frac{1}{N} \sum_i D_c x_k^i + \frac{1}{N} \sum_i \sum_j D_c q_j \frac{\partial \tilde{x}_k^i}{\partial q_j} \right] \\ & + \frac{1}{\lambda} cov \left(D_c q_j \frac{\partial x_j^i}{\partial b}, x_k^i \right) - cov \left(\sum_j t_j \frac{\partial x_j^i}{\partial b}, x_k^i \right). \end{aligned} \tag{4.12}$$

In this formulation, the left-hand side is the same as in the welfarist case and it reflects the aggregate compensated change in the demand of good k. The first two terms in the squared brackets at the right-hand side capture the impacts of tax changes on poverty: the first term is the direct impact of the price change (keeping consumption unaffected) on measured poverty, whereas the second depends on the behavioral shift in comsumption. Multiplied by the minus sign, the former term implies that the consumption of the good should be encouraged, whereas if demand decreases when the prices increase, the latter term actually serves to discourage consumption. Similarly as in equation (2.18), the last two terms at the right arise from aggregation issues and are related to turning the aggregated responses to responses evaluated at the mean demand.

The key lesson to note from the optimal commodity tax rule in the poverty minimization case is that the conventional conditions for uniform commodity tax to be optimal are not valid anymore. The reason is that even if demand was separable from labor supply, the first term at the right still remains in the rule, and its magnitude clearly varies depending on the quantity of good consumed. Thus, income transfers are not sufficient to alleviate poverty when the government aims to minimize poverty that depends on disposable income. The intuition is very simple: commodity tax changes have a direct effect on the purchasing power of the consumer and these depend on the amount consumed. The extent of encouraging the consumption of the goods is the greater the larger is their share of consumption among the consumption bundles of the poor. A formal proof is provided in the Appendix.

In sum, the rule for optimal commodity taxation is changed when we shift from welfare maximization to poverty minimization. The welfaristic rule reflects a fairly straightforward trade-off between efficiency (tax revenue) and equity (distributional impacts). The poverty-minimizing commodity tax rule brings new terms, the interrelations of which are not easy to entangle. It however also takes into account efficiency considerations (tax revenue through indirect labour supply effects) and equity (direct impact of the taxed good on poverty and indirect impact via labour supply effects). Most importantly, the conventional wisdom of when uniform commodity taxation is sufficient fails to hold in the poverty minimization case.

5 Conclusion

This paper examined optimal linear income taxation, public provision of public and private goods and the optimal combination of linear income tax and commodity taxes when the government's aim is to minimize poverty. The linear tax environment was chosen because such taxes are more easily implementable in a developing country context and as they also depend on similar sufficient statistics than more complicated non-linear formulas, this giving rise to similar guidelines for empirical work in the area.

The results show that the linear income tax includes additional components that work towards lowering the marginal tax rate. This result arises from the goal to boost earnings to reduce income poverty. Unlike in the optimal non-linear income tax framework, this lower marginal tax affects all taxpayers in the society. The public good provision in the optimal tax framework under poverty minimization was shown to depend on the relative efficiency of public provision versus income transfers in generating poverty reductions. One particular avenue where public provision is useful is via its potentially beneficial impact on individuals' earnings capacity. Thus, public provision can be desirable even if its direct welfare effects were non-existent. Finally, and perhaps most importantly, poverty minimization as an objective changes completely the conditions under which uniform commodity taxation is optimal. When the government objective is to minimize poverty that depends on disposable income, uniform commodity taxation is unlikely to be ever optimal: this is because the commodity tax changes have first-order effects on consumer's budget via the direct impact on the cost of living, and this direct effect depends on the relative importance of different goods in the overall consumption bundle. Separability in demand coupled with linear Engel curves is not sufficient to guarantee optimality of uniform commodity taxes.

6 Applications

It is our intention to examine various extensions to the analysis in later versions of the paper. These extensions include the possibility that parts of the economy remain informal, and therefore taxes on these activities do not have their conventional impacts; or that state capacity can be limited and therefore public provision could involve additional cost ("leakage") for societies, and numerical illustrations of optimal policies under these scenarios. Here we briefly illustrate the idea of informal sector.

In richer economies, "informality" is usually considered to consist of tax avoidance and evasion behaviour. For example, Piketty, Saez, and Stantcheva (2014) divide the behavioural response to tax rates into a separate avoidance/evasion elasticity along with other elasticities. However, this kind of partial tax compliance can be very different behaviour from that in the developing countries, where individuals might stay completely out of the tax net. For these countries, we consider the informal sector to represent more the lack of opportunity to incorporate oneself into the formal society altogether.

One approach to incorporate this kind of informality into the model is to define a probability κ for an individual to be in the formal sector. Assume that people in the formal sector pay the linear income tax τ in full and receive the income transfer b. Then share $1-\kappa$ remain in the informal sector and don't pay income taxes at all. Assume however that they still receive the income transfer. As the transfer is defined to be a universal lump-sum benefit, everybody receives it regardless of their economic activity. For instance, when receiving the benefit the individual could simply claim the authority that they have not had any income, and assuming there is low administrative capacity, verification would not be feasible. The decision to pay taxes or not is likely to be a function of both the individual income level z^i – the poor are more likely to find it too costly to join the formal sector – and the tax rate τ which alters the relative cost of staying in the formal sector for everyone. The probability is thus defined as $\kappa(\tau, z^i(\tau, b))$, and its derivative with respect to taxes τ as follows: $\kappa\prime = \kappa_\tau + \kappa_z z_\tau$where $\kappa_\tau < 0$, $\kappa_z > 0$ and $z_\tau < 0$ so that the result is $\kappa\prime < 0$. The derivative with respect to the income transfer b in its turn is $\kappa_z z_b < 0$.

The government thus collects tax revenue only from those κ individuals who are in the formal sector, but provides the income transfer to everyone. This necessarily leads to lower benefit size for everyone, since tax revenues are now lower: $\sum_i \kappa\tau z^i < \sum_i \tau z^i$. However, we would expect that this behaviour has a positive effect on poverty: if the poorest are most likely to remain in the informal sector, the income tax does not affect their disposable income level, whereas the richer pay the income taxes (at least more often) which finance the lump-sum transfer to everyone. We thus expect to find the deprivation measure to go down as we apply these considerations. It can thus be optimal for a poverty-minded government to allow some informality at the lower end of the income distribution, at least until tax systems are refined enough to include everyone in the tax net (and possible apply progressive tax rates to favor the poor).[4]

References

Alkire, S., and J. Foster (2011): "Counting and Multidimensional Poverty Measurement," *Journal of Public Economics*, 95, 476–487.

Barrientos, A. (2013): *Social Assistance in Developing Countries*. Cambridge University Press.

Besley, T., and R. Kanbur (1988): "Food Subsidies and poverty Alleviation," *The Economic Journal*, 98, 701–719.

Besley, T., and T. Persson (2013): "Chapter 2 - Taxation and Development," in *Handbook of Public Economics, vol. 5*, ed. by A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez, vol. 5 of *Handbook of Public Economics*, pp. 51 – 110. Elsevier.

Bourguignon, F., and S. Chakravarty (2003): "The Measurement of Multidimensional Poverty," *Journal of Economic Inequality*, 1, 25–49.

Deaton, A. (1979): "Optimally Uniform Commodity Taxes," *Economics Letters*, 2, 357–361.

Diamond, P. (1975): "A Many-Person Ramsey Tax Rule," *Journal of Public Economics*, 4, 335–342.

Dixit, A., and A. Sandmo (1977): "Some Simplified Formulae for Optimal Income Taxation," *Scandinavian Journal of Economics*.

Foster, J., J. Greer, and J. Thorbecke (2010): "The Foster-Greer-Thorbecke (FGT) Poverty Measures: 25 Years Later," *Journal of Economic Inequality*, 8, 491–524.

Gordon, R., and W. Li (2009): "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation," *Journal of Public Economics*, 93(7-8), 855–866.

IFS, and J. Mirrlees (2011): *Tax by design*. Oxford University Press.

Jacquet, L., E. Lehmann, and B. Van der Linden (2013): "Optimal Redistributive Taxation with Both Extensive and Intensive Responses," *Journal of Economic Theory*, 148, 1770–1805.

[4]This set-up of informality and tax collection resembles the intensive-extensive margin analysis of nonlinear taxation, as for example in Jacquet, Lehmann, and Van der Linden (2013).

Kanbur, R., and M. Keen (1989): "Poverty, Incentives and Linear Income Taxation," in *The Economics of Social Security*, ed. by A. Dilnot, and I. Walker, pp. 99–115. Oxford: Clarendon Press.

Kanbur, R., M. Keen, and M. Tuomala (1994): "Optimal Non-Linear Income Taxation for the Alleviation of Income-Poverty," *European Economic Review*, 98, 701–719.

Kanbur, R., J. Pirttila, and M. Tuomala (2006): "Non-Welfarist Optimal Taxation and Behavioural Public Economics," *Journal of Economic Surveys*, 20, 849–868.

Keen, M. (2012): "Taxation and Development: Again," IMF Working Papers 12/220, International Monetary Fund.

Mirrlees, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(114), 175–208.

Piketty, T., and E. Saez (2013): "Chapter 7 - Optimal Labor Income Taxation," in *Handbook of Public Economics, vol. 5*, ed. by A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez, vol. 5 of *Handbook of Public Economics*, pp. 391 – 474. Elsevier.

Piketty, T., E. Saez, and S. Stantcheva (2014): "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 6, 230–271.

Pirttila, J., and M. Tuomala (2004): "Poverty Alleviation and Tax Policy," *European Economic Review*, 48, 1075–1090.

Tuomala, M. (1985): "Simplified Formulae for Optimal Linear Income Taxation," *Scandinavian Journal of Economics*, 87, 668–672.

A Appendix 1: Proof of nonuniform commodity taxation optimality

We demonstrate formally how uniform commodity taxation is not optimal in the case of poverty minimization. To see this, rewrite first the FOC with respect to b (equation (4.8)) as

$$\frac{1-\frac{1}{N}\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b}}{\frac{1}{N}\sum_i D_c \sum_j q_j \frac{\partial x_j^i}{\partial b}} = \frac{1}{\lambda}. \tag{A.1}$$

Next, rewriting the FOC for q_k (equation (4.10)) yields

$$\frac{1}{N}\sum_i \sum_j t_j \frac{\partial \tilde{x_k}^i}{\partial q_j} = -\frac{1}{\lambda}\frac{1}{N}\sum_i D_c \frac{x_k^i}{N} - \frac{1}{\lambda}\frac{1}{N}\sum_i D_c \sum_j q_j \frac{\partial \tilde{x_k}^i}{\partial q_j} + \frac{1}{\lambda}\frac{1}{N}\sum_i D_c \sum_j q_j \frac{\partial x_j^i}{\partial b} x_k^i + \frac{1}{N}\left(\sum_j t_j \frac{\partial x_j^i}{\partial b} - 1\right)\sum_i x_k^i \quad \text{(A.2)}$$

Substituting for $\frac{1}{\lambda}$ from (A.1) in the first term at the lower row of equation (A.2) gives

$$\frac{1}{N}\sum_i \sum_j t_j \frac{\partial \tilde{x_k}^i}{\partial q_j} = -\frac{1}{\lambda}\frac{1}{N}\sum_i D_c \frac{x_k^i}{N} - \frac{1}{\lambda}\frac{1}{N}\sum_i D_c \sum_j q_j \frac{\partial \tilde{x_k}^i}{\partial q_j} + \frac{1 - \frac{1}{N}\sum_i \sum_j t_j \frac{\partial x_j^i}{\partial b}}{\sum_i D_c \sum_j q_j \frac{\partial x_j^i}{\partial b}} \sum_i D_c x_k^i \sum_j q_j \frac{\partial x_j^i}{\partial b} + \frac{1}{N}\sum_i x_k^i \left(\sum_j t_j \frac{\partial x_j^i}{\partial b} - 1\right). \quad \text{(A.3)}$$

Following Deaton (1979, p. 359-360), when preferences are separable and Engel curves are linear, demand is written as $x_j^i = \delta_j^i(q) + \theta_j(q)c^i$, hence the derivative of demand with respect to disposable income c or transfer b is $\theta_j(q)$, i.e. independent of the person i. By writing out explicitly the solution that the derivative of demand w.r.t b is independent of i and write $\frac{\partial x_j^i}{\partial b} = \theta_j(q)$ we have:

$$\frac{1}{N}\sum_i \sum_j t_j \frac{\partial \tilde{x_k}^i}{\partial q_j} = -\frac{1}{\lambda}\frac{1}{N}\sum_i D_c \frac{x_k^i}{N} - \frac{1}{\lambda}\frac{1}{N}\sum_i D_c \sum_j q_j \frac{\partial \tilde{x_k}^i}{\partial q_j} + \frac{1 - \frac{1}{N}\sum_i \sum_j t_j \beta_j(q)}{\sum_i D_c \left(\sum_j q_j \beta_j(q)\right)} \sum_i D_c x_k^i \left(\sum_j q_j \beta_j(q)\right) + \frac{1}{N}\sum_i x_k^i \left(\sum_j t_j \beta_j(q) - 1\right) \quad \text{(A.4)}$$

where in the second row we can cancel out the $\sum_j q_j \beta_j(q)$ terms and rewrite $\sum_i \sum_j t_j \beta_j(q) = N\sum_j t_j \beta_j(q)$ in the numerator because the term is independent over i:

$$
\begin{aligned}
\frac{1}{N}\sum_i\sum_j t_j\frac{\partial\tilde{x_k}^i}{\partial q_j} \quad = \quad & -\frac{1}{\lambda}\frac{1}{N}\sum_i D_c\frac{x_k^i}{N}-\frac{1}{\lambda}\frac{1}{N}\sum_i D_c\sum_j q_j\frac{\partial\tilde{x_k}^i}{\partial q_j} \\
& +\frac{1-\sum_j t_j\beta_j(q)}{\sum_i D_c}\sum_i D_c x_k^i+\frac{1}{N}\sum_i x_k^i\left(\sum_j t_j\beta_j(q)-1\right) \quad \text{(A.5)}
\end{aligned}
$$

Note next that due to homegeneity of degree 0 of compensated demand, $\sum_j q_j\frac{\partial\tilde{x_k}^i}{\partial q_j}+w_i\frac{\partial\tilde{x_k}^i}{\partial w_i}=0$. This, together with the observation that if a uniform commodity tax t was a solution to a problem at hand, this would mean that the left-hand side of (A.2) could be written as $-\frac{t}{N}\sum_i w_i\frac{\partial\tilde{x_k}^i}{\partial w}$. Because of separability, the substitution response is linked to the full income derivative, so that $\frac{\partial\tilde{x_k}^i}{\partial w}=\phi^i\theta_j(q)$. Because of these arguments, (A.2) becomes

$$
\begin{aligned}
-\frac{t}{N}\theta_j(q)\sum_i w_i\phi^i \quad = \quad & -\frac{1}{\lambda}\frac{1}{N}\sum_i D_c\frac{x_k^i}{N}-\frac{1}{\lambda}\frac{1}{N}\theta_j(q)\sum_i D_c w_i\phi^i \\
& +\frac{1-t\sum_j\theta_j(q)}{\sum_i D_c}\sum_i D_c x_k^i+\frac{1}{N}\sum_i x_k^i\left(t\sum_j\theta_j(q)-1\right). \quad \text{(A.6)}
\end{aligned}
$$

Note that terms incorporatin $\theta_j(q)$ cannot be cancelled out from the equation so the result remains dependent on j. In addition, even if the terms were cancelled, the term $\sum_i D_c\frac{x_k^i}{N}$ still depends on j. This shows that uniform commodity taxation is not optimal when the objective function of the government is to minimize poverty.

B Appendix 2: Measuring poverty with the FGT poverty index

B.1 Linear income tax

One of the most popular poverty measures is the P_α category developed by Foster, Greer and Thorbecke. It is usually written in the form of $P_\alpha=\int_0^z\left(\frac{z-y}{z}\right)^\alpha f(y)\,\mathrm{d}(y)$ where z is the poverty line and y is income. Following Kanbur and Keen (1989), who define the poverty index in terms of disposable income, and using our notation, this becomes: $P_\alpha=\int_0^{\bar{c}}\left(\frac{\bar{c}-c^i}{\bar{c}}\right)^\alpha\mathrm{d}\nu(i)$, where c^i is disposable income, in the linear tax case $c^i=(1-\tau)z^i+b=(1-\tau)z^i+\tau Z(1-\tau)-R$. In the Piketty and Saez model, we can use this specification of the functional form to define the derivative $D_c=-\frac{\alpha}{\bar{c}}\left(\frac{\bar{c}-c^i}{\bar{c}}\right)^{\alpha-1}$ (note that $D_c<0$ as long as $c^i<\bar{c}$). We can follow the same steps to arrive at the optimal tax rate $\tau^*=\frac{1-\bar{\beta}-\bar{\beta}^e}{1-\bar{\beta}-\bar{\beta}^e+e}$ where now

$$
\beta^i \quad = \quad \frac{D_c}{\int_0^{\bar{c}}D_c\,\mathrm{d}\nu(i)}=\frac{-\frac{\alpha}{\bar{c}}\left(\frac{\bar{c}-c^i}{\bar{c}}\right)^{\alpha-1}}{\int_0^{\bar{c}}-\frac{\alpha}{\bar{c}}\left(\frac{\bar{c}-c^i}{\bar{c}}\right)^{\alpha-1}\mathrm{d}\nu(i)}=\frac{\left(\frac{\bar{c}-c^i}{\bar{c}}\right)^{\alpha-1}}{\int_0^{\bar{c}}\left(\frac{\bar{c}-c^i}{\bar{c}}\right)^{\alpha-1}\mathrm{d}\nu(i)}
$$

and consequently $\bar{\beta} = \frac{\int_0^{\bar{c}} \beta^i z^i \, \mathrm{d}\nu(i)}{Z}$ and $\bar{\beta}^e = \frac{\int_0^{\bar{c}} \beta^i z^i e_c^i \, \mathrm{d}\nu(i)}{Z}$ as before. Everything else stays exactly the same as in the calculations of section 2.2.1. Also in the case of Tuomala's and Dixit and Sandmo's models, the results stay the same, and we can plug in the explicit definition for D_c, the derivative of the poverty measure with respect to disposable income, into the results.

B.2 Public good provision

Employing the FGT poverty measure in the context of public good provision for poverty reduction is a bit more complicated than in the linear tax case. In section 3.2 the government's objective function was defined as $min\, P = \int D\left(x^i, G, \bar{x}, \bar{G}\right) \mathrm{d}\nu(i)$, that is, deprivation was measured both as deprivation in private consumption (i.e. disposable income) as well as with respect to the public good. But the FGT measure is a unidimensional measure, measuring deprivation with respect to one dimension only (e.g. disposable income). If one wants to consider publicly offered goods such as education as separate from private consumption, a multidimensional FGT measure is needed. Multidimensionality however entails a difficult question of determining when a person can be determined as deprived. (Note that in section 4.2, which covered commodity taxation, we escaped this issue since we defined deprivation in terms of total consumption only, not making distinctions between different commodites.)

Bourguignon and Chakravarty (2003) provide a multidimensional extension of the FGT measure, according to which a person is poor if she is deprived in at least one dimension. A simple example of such an extension of the FGT is

$$P_\theta = \frac{1}{n} \sum_{j=1}^{m} \sum_{i \in S_j} a_j \left(\frac{z_j - x_{ij}}{z_j} \right)^{\theta_j}$$

where θ_j and a_j are weights given to dimension j, and S_j is the group of people who are poor in dimension j. Alkire and Foster (2011) for their part provide a measure which uses a weighted count of of dimensions in which the person is deprived to determine whether she is poor. An aspect of this is also whether the goods under consideration are complements or substitutes.[5]

Besley and Kanbur (1988), who consider the poverty impacts of food subsidies, employ the unidimensional FGT measure but define deprivation in terms of equivalent income: $P_\alpha = \int_0^z \left(\frac{z_E - y_E}{z_E} \right)^\alpha f(y) \, \mathrm{d}(y)$, where y_E is equivalent income, defined implicitly from $V(p, y_E) = V(q, y)$, and z_E is the poverty line corresponding to equivalent income. But given our aim of defining optimal policy in terms of poverty reduction, irrespective of individual welfare, the use of equivalent income is problematic as it forces the solution to be such that, by definition, individuals are kept as well off as before. Indeed, Kanbur and Keen (1989) opt out of the equivalent income approach as well: "The equivalent income approach empbodies the

[5]See Foster, Greer, and Thorbecke (2010, p.504-5) for a brief overview of multidimensional FGT extensions that allow the inclusions of dimensions such as health, education, and nutrition in addition to other consumption.

welfarism characteristic of the optimal tax literature. Indeed, it amounts to little more than imposing additional structure on a social welfare function of the usual kind. For this reason, we shall explore this approach no further." (Kanbur and Keen (1989), p.102-103) The equivalent income approach is of course useful for the purpose of evaluating the impacts on poverty of the price changes that public provision can cause, as in Besley and Kanbur (1988). We are not pursuing this, however. Pirttila and Tuomala (2004) follow a similar approach to allow for several goods in the poverty measure. For them, deprivation is measured as $D\left(z, y\left(q, w\right)\right)$ where $z^h = s_x x^* - s_L^h L^*$and $y^h(q, w^h) = s_x x(q, w^h) - s_L^h L(q, w^h)$. This approach requires determining shadow prices for consumption and leisure in order to construct a reference bundle respective to which deprivation can be measured, but there is no clear guideline to the choice of the shadow prices.

There are thus several approaches of using an FGT-type of poverty measure in the determination of optimal provision of public goods for poverty reduction, but it is not clear which one is the 'best' one. The Bourguignon-Chakravarty approach seems the most promising one (it is very simple and intuitive). Defining $x_{i1} = x_i$ as private consumption, $z_1 = \bar{x}$, $x_{i2} = G$ as the amount of public good, and $z_2 = \bar{G}$ would give us $P_\theta = \frac{1}{n}\sum_{i\in S_j}\left(a_1\left(\frac{\bar{x}-x^i}{\bar{x}}\right)^{\theta_1} + a_2\left(\frac{\bar{G}-G}{\bar{G}}\right)^{\theta_2}\right)$. Using this measure, $D_x = -\frac{\theta_1 a_1}{\bar{x}}\left(\frac{\bar{x}-x^i}{\bar{x}}\right)^{\theta_1-1}$ and $D_G = -\frac{\theta_2 a_2}{\bar{G}}\left(\frac{\bar{G}-G}{\bar{G}}\right)^{\theta_2-1}$. These can then be inserted to the public provision rules. For example, (3.7) becomes

$$\frac{\int\left(\frac{\theta_2 a_2}{\bar{G}}\left(\frac{\bar{G}-G}{\bar{G}}\right)^{\theta_2-1} + \frac{\theta_1 a_1}{\bar{x}}\left(\frac{\bar{x}-x^i}{\bar{x}}\right)^{\theta_1-1}(1-\tau)\frac{\partial z^i}{\partial G}\right)\mathrm{d}\nu(i)}{\int\frac{\theta_1 a_1}{\bar{x}}\left(\frac{\bar{x}-x^i}{\bar{x}}\right)^{\theta_1-1}\mathrm{d}\nu(i)} = p - \tau\frac{dZ}{dG}$$

and (3.11) becomes

$$\frac{\sum\frac{\theta_2 a_2}{\bar{G}}\left(\frac{\bar{G}-G}{\bar{G}}\right)^{\theta_2-1} + \sum\frac{\theta_1 a_1}{\bar{x}}\left(\frac{\bar{x}-x^i}{\bar{x}}\right)^{\theta_1-1} a z_G^i}{\sum\frac{\theta_1 a_1}{\bar{x}}\left(\frac{\bar{x}-x^i}{\bar{x}}\right)^{\theta_1-1}\left(1+a z_b^i\right)} = \frac{p-(1-a)\sum z_G^i}{1-(1-a)\sum z_b^i},$$

from where it can be seen that the relative efficiency of the public good versus cash transfers on reducing poverty can be directly traced back to the magnitudes of θ_1 and θ_2.

B.3 Commodity taxation

In the case of commodity taxes, we run into the same issues regarding deprivation measurement as with public goods. However, in section 4.2 deprivation was measured only in terms of disposable income, c. Employing the FGT poverty measure is thus as simple as in the linear income tax case, we simply need to define $D = P^\alpha$ and thus $D_c = -\frac{\alpha}{\bar{c}}\left(\frac{\bar{c}-c^i}{\bar{c}}\right)^{\alpha-1}$ in equation (4.12). Potentially the government might also consider weighting different goods according to their importance to measured poverty.