Discussion Paper No. 14-002

Fair and Efficient Taxation Under Partial Control

Erwin Ooghe and Andreas Peichl

Zentrum für Europäische Wirtschaftsforschung GmbH

Centre for European Economic Research

Discussion Paper No. 14-002

Fair and Efficient Taxation Under Partial Control

Erwin Ooghe and Andreas Peichl

Fair and efficient taxation under partial control*

Erwin Ooghe · Andreas Peichl†

19th December 2013

Abstract

We study fair and efficient tax-benefit schemes based on income and non-income factors under partial control. Partial control means that each factor is a specific mixture of unobserved ability (randomly drawn by nature) and effort (chosen by individuals who differ in tastes). Factors differ in the degree of control, ranging from no control (if only ability matters) to full control (if only effort matters). Fairness requires to compensate individuals for differences in well-being caused by differences in abilities, while at the same time preserving well-being differences caused by taste differences. We discuss first the general properties of fair and efficient tax-benefit schemes. Next, we study two special cases—income taxation and tagging—in detail. Finally, we derive testable conditions for the general case and discuss the empirical implementation.

JEL codes: D6, H2, I3

Keywords: fairness, redistribution, taxation, tagging, equality of opportunity.

*We would like to thank Martin Cripps, the editor, and two anonymous referees as well as Spencer Bastani, Koen Decancq, André Decoster, Clemens Fuest, Laura Kalambokidis, Dirk Neumann, Andrew Oswald, Jukka Pirttilä, Jim Poterba, John Roemer, Erik Schokkaert, Alain Trannoy, Andreas Wagener, participants at seminars in Barcelona (UAB), Bonn (IZA), Louvain-La-Neuve (CORE), Mannheim (ZEW), Munich (CESifo), and participants at conferences in Chicago (NTA), Marseille (LAGV), Sankt Gallen (IARIW), and Uppsala (IIPF) for helpful comments and suggestions.

†Erwin Ooghe, Department of Economics (KULeuven) and IZA; e-mail to erwin.ooghe@kuleuven.be. Andreas Peichl, ZEW, University of Mannheim and IZA; e-mail to peichl@zew.de.

1 Introduction

Economists traditionally assume that individuals are motivated only by their material self-interest. But experiments systematically reject the pure self-interest hypothesis; see, e.g., Fehr and Schmidt (2006) for a survey. Other considerations, like fairness, do play a role. If earnings are a combination of brute luck (drawn by nature) and effort (chosen by the individual), then people are willing to compensate others for unlucky draws by nature, but also allow them to enjoy the fruits of their effort. Empirical evidence shows that the more income is determined by luck, the more redistribution is preferred. Konow (2003), Alesina and Giuliano (2010), and Gaertner and Schokkaert (2011) provide overviews based on laboratory experiments, social survey data, and structured questionnaires.

Fairness considerations have been introduced in political economy models. Alesina et al. (2001) show that different beliefs about the importance of luck for income acquisition can help explain the divergence in redistribution levels in different democratic societies. The political economy models of Piketty (1995), Alesina and Angeletos (2005), Bénabou and Tirole (2006), and Alesina et al. (2012) lead to multiple states such that stronger beliefs in the role of effort coincide with lower levels of redistribution.[1]

Similar notions of fairness have been introduced in the so-called fair income tax literature. Some authors start from a specific fairness notion, derive the corresponding social ordering, and characterise or simulate optimal income tax schemes; see, e.g., Roemer et al. (2003), Schokkaert et al. (2004), Fleurbaey and Maniquet (2006, 2007), Luttens and Ooghe (2007), Jacquet and Van de gaer (2011), and Aaberge and Colombino (2012). Other authors study the consequences of introducing preference heterogeneity directly in the optimal (utilitarian) income tax literature; see, Boadway et al. (2002), Kaplow (2008), Choné and Laroque (2010), and Lockwood and Weinzierl (2012). Contrary to Mirrlees (1971), negative marginal income taxes—subsidies to the hard-working poor—may be optimal.

Political economy models and fair income tax models traditionally focus on earnings only. There exist, however, different theoretical reasons to include also non-income information in the tax base. If externalities exist, then there is a role for government to subsidise or tax these activities à la Pigou (1920) to restore efficiency. If there exist tags—observable, usually exogenous factors that correlate with unobserved abilities or

[1] See also the comment on Alesina and Angeletos (2005) by Di Tella and Dubra (2013) and the reply by Alesina et al. (2013).

tastes—then differentiating the tax-benefit system on the basis of these tags (sometimes called tagging) can also enhance efficiency; see Akerlof (1978) for his seminal contribution.[2] The optimal income tax treatment of family size and couples also received considerable attention; see, e.g., Mirrlees (1972) and Boskin and Sheshinski (1983) for initial contributions.[3]

In this paper, we study the design of fair and efficient tax-benefit schemes based on income and non-income factors under partial control. We preview the core ingredients.

Individuals differ in unobserved abilities and tastes.[4] Taste differences bring the question of fairness—which inequalities are justifiable and which are not—to the fore. Fleurbaey and Maniquet (2006) propose to keep individuals responsible for their tastes, but to compensate them for differences in their abilities. Responsibility for tastes demands that the laisser-faire should result if all individuals have the same ability. In this case, differences in outcomes can only be caused by differences in tastes for which they were kept responsible. Compensation for abilities requires to approve of transfers from better off to worse off if these differences in income are caused only by differences in abilities (i.e., if they have the same preferences and exert the same effort). We use a classical welfare function—a sum of transformed utilities—that satisfies the Pareto principle, compensation, and responsibility.

Besides income, we also model non-income factors. Both income and non-income factors are modelled as a convex combination of ability (drawn by nature) and effort (chosen by individuals who differ in tastes). The weight defines the degree of control. For some factors, think of an inborn handicap, the degree of control is zero, while other factors, think of earnings or family composition, the degree of control is positive and partial control applies.

The complexity of the resulting multidimensional screening exercise forces us to simplify several aspects of the model to keep analytical tractability. Besides a linear production technology under partial control, we assume quasi-linear preferences (defined over

[2] Tagging has also been analysed by, among others, Immonen et al. (1998) and Salanié (2002, 2003). While these authors do not have a specific tag in mind, Blomquist and Micheletto (2008), Bastani et al. (2013), and Weinzierl (2011) consider age tags, Mankiw and Weinzierl (2010) study height, and Cremer et al. (2010) and Alesina et al. (2011) focus on gender.

[3] See, e.g., Cremer et al. (2003, 2012), Schroyen (2003), Brett (2007), Kleven et al. (2009) and Immervoll et al. (2011) for a recent state-of-the-art.

[4] The standard optimal income tax literature is traditionally based on ability heterogeneity only; see recent surveys by Mankiw et al. (2009), Diamond and Saez (2011), Boadway (2012), and Piketty and Saez (2013).

income, non-income factors, and effort), independent multivariate normal distributions for abilities and tastes, and linear tax rates for the income and non-income factors.

Our results show that, in general, optimal tax rates balance the marginal efficiency cost of taxation caused by tax distortions against the marginal net fairness benefit of taxation. The latter is the difference between the marginal compensation benefit and the marginal responsibility cost of taxation. Higher taxes reduce outcome differences between individuals with the same tastes, but different abilities—a good thing—reflected by the marginal compensation benefit of taxation. Yet, higher taxes also reduce the outcome differences between individuals with the same abilities, but different tastes—a bad thing—captured by the marginal responsibility cost of taxation.

We also study two special cases in detail. In case only income is included in the model, we show, among other things, that the optimal income tax negatively depends on the degree of control over income. The tax must also increase with ability heterogeneity, while it has to decrease with taste heterogeneity. If we add a tag to the model—an observable non-controllable non-income factor, say, an inborn handicap—, then the optimal tax on the tag depends on the correlation between the tag and the unobserved ability to earn income. Introducing taste heterogeneity and tagging lowers the optimal tax on income.

Finally, we show how the theory can be tested empirically. We derive testable conditions for the tax rates on non-controllable factors in the general case. They turn out to be equal to the sum of the direct effects of non-controllable factors on well-being augmented by their indirect effects via the correlation with partially controllable factors. We also discuss how the theory can be empirically tested using happiness data. We show that in this case, the theory obtains a simple structure that resembles equality of opportunity regressions.

The remainder of the paper is structured as follows. Section 2 describes the model. Section 3 discusses the main result. Section 4 and 5 study two special cases—income taxation and tagging—in detail. Section 6 returns to the general case and derives testable conditions. Section 7 concludes.

2 The model

We define the basic building blocks—preferences and constraints—at the individual and the societal level.[5]

2.1 Individual preferences and constraints

Individual utility $U(c, x, e)$ is a function of consumption $c \in \mathbb{R}$, non-income factors $x = (x_1, x_2, \ldots, x_J) \in \mathbb{R}^J$, and effort $e = (e_0, e_1, \ldots, e_J) \in \mathbb{R}^{J+1}$. Consumption c equals gross income y minus taxes $\tau(y, x)$, a function of gross income and non-income factors. A production function $f : \mathbb{R}^{J+1} \to \mathbb{R}^{J+1}$ maps effort e into gross income and non-income factors (y, x).

The utility maximising problem of an individual can be summarised as

$$\max_e \ U(c, x, e) \text{ subject to } c \leq y - \tau(y, x) \text{ and } (y, x) = f(e).$$

To keep the model tractable, we make the following simplifying assumptions.

1. **Quasi-linear preferences.** Utility is equal to consumption plus the value of non-income factors minus the cost of effort, or

   $$U(c, x, e) = c + g(x) - h(e).$$

   This specification is quasi-linear in consumption, as in Diamond (1998), and excludes therefore income effects.[6] In addition, the value function of non-income factors g and the cost function of effort h have a flexible parameterisation, more precisely,

   $$\begin{aligned} g(x) &= \textstyle\sum_{j=1}^{J} \beta_j x_j, \\ h(e) &= \textstyle\sum_{j=0}^{J} \dfrac{\delta_j}{\exp \gamma_j} \exp(\dfrac{e_j}{\delta_j}). \end{aligned}$$

   The vector β captures the trade-off between income and non-income factors in a simple linear way, so it extends quasi-linearity to the non-income factors. Without loss of generality we assume the non-income factors to be desirable, i.e., $\beta \in \mathbb{R}^J_{++}$. The vector $\gamma \in \mathbb{R}^{J+1}$ is a taste vector, defining the disutility of effort. Higher values

[5] We present an additive version of the model here. The appendix in Ooghe and Peichl (2010) contains a multiplicative variant with the same results.

[6] Zero income effects are often not falsified by the data; see, e.g., the discussion in Diamond and Saez (2011) or Piketty and Saez (2013) as well as the results in Bargain et al. (2014).

for γ correspond with lower disutilities of effort, thus more ambitious individuals. The vector $\delta \in \mathbb{R}_{++}^{J+1}$ is an elasticity vector that controls the convexity of h. Higher values for δ correspond with more elastic responses to effort and thus a higher cost of taxation.

2. **Linear production.** Gross income y and the non-income factors in x are under partial control, i.e., they are each a convex combination of effort and ability. We define

$$\begin{aligned} y &= \alpha_0 e_0 + (1-\alpha_0)\,\theta_0, \\ x_j &= \alpha_j e_j + (1-\alpha_j)\,\theta_j \quad \text{for} \quad j = 1, 2, \ldots, J, \end{aligned}$$

with $\alpha \in (0,1)^{J+1}$ collecting the weights and $\theta \in \mathbb{R}^{J+1}$ the abilities. The weights define the degree of control of a factor, ranging from no control ($\alpha_j \to 0$; the factor is ability only) over partial control ($0 < \alpha_j < 1$) to full control ($\alpha_j \to 1$; the factor is effort only).

3. **Ability and taste heterogeneity.** Heterogeneity in abilities is modelled via heterogeneity in the vector θ. In particular, we assume a multivariate normal distribution, with $\mu^\theta = (\mu_0^\theta, \mu_1^\theta, \ldots, \mu_J^\theta)$ the vector of means and $\Sigma^\theta = [\sigma_{jk}^\theta]$ the variance-covariance matrix, with $\sigma_{jk}^\theta = E\{(\theta_j - \mu_j^\theta)(\theta_k - \mu_k^\theta)\}$.[7] To model heterogeneity in tastes, we assume that the taste vector γ also follows a multivariate normal distribution, fully described by μ^γ and Σ^γ. Note that abilities and tastes are assumed to be independently distributed.[8] This assumption avoids the philosophical problem whether we can keep individuals responsible for their tastes, if the latter correlate with abilities. Still, independence between abilities and tastes does not exclude that income and the non-income factors are correlated in complex ways. We will come back to this issue in detail in section 6.

Individuals know their abilities and tastes when choosing effort.[9] Let $e^*(\tau, \theta, \gamma)$ be the utility maximising effort choice and $c^*(\tau, \theta, \gamma)$, $y^*(\tau, \theta, \gamma)$, and $x^*(\tau, \theta, \gamma)$ the corresponding (net and gross) income and the non-income factors.

[7] We exclude perfect correlation, so $(\sigma_{jk}^\theta)^2 < \sigma_{jj}^\theta \sigma_{kk}^\theta$ for $j \neq k$.

[8] Whereas abilities θ and tastes γ are heterogeneous, the degrees of control α, the parameters in β, and the elasticities δ are the same for all individuals.

[9] The effort choice would remain the same if individuals were expected utility maximisers that know their tastes, but only the distribution of abilities. Abilities can thus also be interpreted as risks.

2.2 Social preference and constraint

The social preference and its owner, the fictitious social planner, are a proxy for a more complex political voting model.[10] The problem of the planner is to choose a tax scheme τ to maximise welfare subject to a budget constraint. Let R_0 denote an exogenous (per-capita) revenue requirement; the social planner's problem is

$$\max_{\tau} \; W(\tau) \text{ subject to } R(\tau) \geq R_0,$$

with W the welfare function and R the revenue function. The revenue function measures the average tax revenue, so $R(\tau) = \int_\theta \int_\gamma \tau(y^*(\tau,\theta,\gamma), x^*(\tau,\theta,\gamma))dF(\theta)\, dG(\gamma)$, with F and G the distribution functions of abilities and tastes. We make again some additional simplifying assumptions.

1. **Welfare.** Welfare is a sum of transformed well-being levels, more precisely

   $$W(\tau) = \phi^{-1}[\textstyle\int_\theta \int_\gamma \phi(v(\tau,\theta,\gamma))dF(\theta)dG(\gamma)],$$

   with ϕ a strictly increasing transformation function and $v(\tau,\theta,\gamma)$ the well-being level of an individual (explained in the next item). The transformation function ϕ is chosen to be exponential, i.e., $\phi : x \mapsto \exp(-rx)$, with $r > 0$ the degree of inequality aversion. This leads to a Kolm-Pollak welfare function in between utilitarianism ($r \to 0$) and maximin ($r \to +\infty$).

2. **Well-being.** Inspired by Fleurbaey and Maniquet (2006)—and for reasons that will be explained below—the well-being function v is chosen to be a specific cardinalisation of the indirect utility function V. Indirect utility is defined as

   $$V(\tau,\theta,\gamma) = U(c^*(\tau,\theta,\gamma), x^*(\tau,\theta,\gamma), e^*(\tau,\theta,\gamma)),$$

   and well-being $\widehat{v} = v(\tau,\theta,\gamma)$ is implicitly defined by the equation

   $$V(\tau,\theta,\gamma) = V(R_0, (\widehat{v}, \widehat{v}, \ldots, \widehat{v}), \gamma).$$

   In words, $\widehat{v} = v(\tau,\theta,\gamma)$ is the ability level that makes an individual indifferent between his actual situation, with utility level $V(\tau,\theta,\gamma)$, and the following hypothetical situation:[11]

[10] Mueller (2003) shows the equivalence between a planner with a weighted utilitarian welfare function and a probabilistic voting model with two candidates that compete for votes. In the current context, Alesina et al. (2012) and Alesina et al. (2013) use probabilistic voting models to analyse fairness and redistribution.

[11] In the appendix we show that $\widehat{v}$ is well-defined and unique under the assumptions made.

(a) the tax is lump-sum and satisfies the revenue constraint, so, $\tau = R_0$,

(b) the individual has a hypothetical ability type equal to $(\widehat{v}, \widehat{v}, \ldots, \widehat{v})$,

(c) the individual keeps his own preferences parameterised by γ.

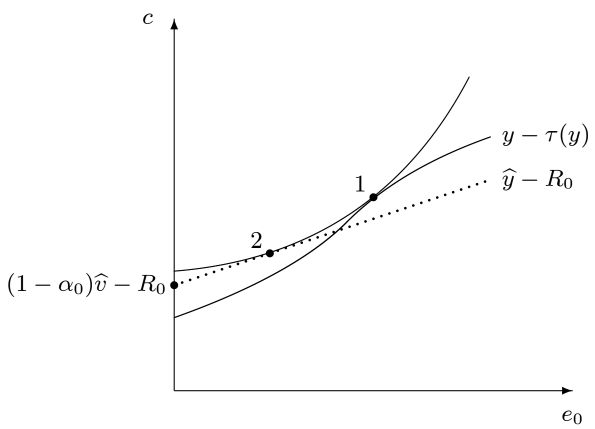
3. **Linear taxation.** Taxation is linear, i.e.,

$$\tau(y, x) = T + t_0 y + \textstyle\sum_{j=1}^{J} t_j x_j,$$

with $T \in \mathbb{R}$ the demogrant and $t = (t_0, t_1, \ldots, t_J) \in \mathbb{R}^{J+1}$ the tax rates that apply to income and the non-income factors. Linearity is restrictive, but it is nonetheless a good approximation of existing tax-benefit schemes.[12]

We illustrate and justify the construction of well-being as a specific cardinalisation of indirect utility. Figure 1 illustrates the construction of well-being in case only income matters; the tax is a general function of income and denoted by $\tau(y)$ in the absence of non-income factors.

Figure 1: Well-being in the absence of non-income factors



Bundle 1 is the actual choice from the budget set $c \leq y - \tau(y)$ for some arbitrary tax scheme τ. Bundle 2 is the hypothetical choice from the budget set $c \leq \widehat{y} - R_0$, with $\widehat{y}$ the hypothetical gross income defined as $\widehat{y} = \alpha_0 e_0 + (1 - \alpha_0)\widehat{v}$, and $\widehat{v}$ chosen such that the individual is indifferent between the actual and hypothetical choice. The well-being level

[12] The total variation in net tax payments in Europe and the United States can be linearly explained by non-income factors (49% on average) and income (30%), while higher-order terms for income do not seem to play an important role (5%). The remaining part is either unexplained (12%) or reflects covariances between the observed characteristics (4%); see Ooghe and Peichl (2010) for details.

$\widehat{v}$ can be read—up to an affine transformation—at the intersection of the hypothetical budget line and the vertical axis.

The social ranking of tax schemes results as an inextricable combination of a specific welfare and well-being function. In line with Fleurbaey and Maniquet (2006), it satisfies the Pareto principle and two fairness principles, responsibility and compensation. We provide an informal discussion of the properties here; a formal statement can be found in the appendix.

The Pareto principle guarantees that the social planner will select only Pareto efficient tax schemes, i.e., in the optimal tax scheme no one can be made strictly better off without making some other people strictly worse off. Responsibility and compensation implement the fairness idea that keeps individuals responsible for their tastes, but compensates them for differences in abilities. Responsibility requires the laisser-faire ($\tau = R_0$) to be optimal in case all individuals have the same abilities. In this peculiar case, differences between individuals are entirely driven by differences in tastes; and if individuals are responsible for their tastes, then there is indeed no reason for redistribution. The compensation principle approves of income transfers from the better off to the worse off in case two individuals have the same tastes. In this specific case, differences in outcomes between both individuals can be traced back to differences in their abilities only; and if they are not responsible for their abilities, then income redistribution is justified.

3 Main result

Proposition 1 characterises the general solution; all proofs can be found in the appendix.

Proposition 1. The optimal tax rate vector $t^* = (t_0^*, t_1^*, \ldots, t_J^*)$ must satisfy the first-order conditions

$$\zeta \times \underbrace{\alpha_j \delta_j \frac{t_j}{\beta_j - t_j}}_{\text{efficiency cost}} = r \times \{\underbrace{(1-\alpha_j)\textstyle\sum_{k=0}^{J}(\beta_k - t_k)(1-\alpha_k)\sigma_{kj}^{\theta}}_{\text{compensation benefit}} - \underbrace{\alpha_j \delta_j \textstyle\sum_{k=0}^{J} t_k \alpha_k \delta_k \sigma_{kj}^{\gamma}}_{\text{responsibility cost}}\},$$

for $j = 0, 1, \ldots, J$, with $\beta_0 \equiv 1$ and $\zeta \equiv \sum_{j=0}^{J}(1-\alpha_j)\beta_j > 0$.

The left-hand side of each first-order condition is—up to the scale factor ζ—the marginal efficiency cost of taxation caused by tax distortions. The inequality aversion parameter r plays indeed no role here. The marginal efficiency cost of taxation approaches zero if the taxed factor cannot be changed by effort ($\alpha_j \to 0$), if the factor is inelastically provided ($\delta_j \to 0$), or if the tax rate is equal to zero ($t_j \to 0$). It becomes infinitely large if the tax rate confiscates the complete value of a factor ($t_j \to \beta_j$). As a consequence we have $t_j < \beta_j$ at the optimum, for all $j = 0, 1, \ldots, J$.

The right-hand side is the net marginal fairness benefit of taxation (the term between curly brackets) weighted by the inequality aversion parameter r. If society cares only about efficiency ($r \to 0$), then taxation only causes distortions and the optimal linear tax scheme, denoted (T^*, t^*), must coincide with the laisser-faire tax scheme $(R_0, 0)$.

The *net* marginal fairness benefit of taxation is equal to the marginal compensation benefit of taxation minus the marginal responsibility cost. The first term between curly brackets is the marginal compensation benefit of taxation. It reflects the fact that higher taxes reduce outcome differences between people with the same tastes, but different abilities. Viewed in this way, taxation compensates individuals for ability differences, and thus increases welfare. The marginal compensation benefit depends on the degrees of control and on the variance-covariance structure of abilities.

The second term between curly brackets is the marginal responsibility cost of taxation. Higher taxes also reduce outcome differences between people with the same abilities, but different tastes. Viewed from this angle, taxation goes against responsibility and thus decreases welfare. The responsibility term therefore enters as a cost. The marginal responsibility cost depends on the degrees of control, the elasticities, and the variance-covariance structure of tastes.

The marginal compensation benefit of taxation is equal to zero if there is no ability heterogeneity ($\Sigma^\theta \to 0$). If everyone has the same ability, then taxation only causes costs—efficiency and responsibility costs—and the laisser-faire will be optimal, as required indeed by the responsibility principle. The marginal responsibility cost of taxation becomes zero if there is no taste heterogeneity ($\Sigma^\gamma \to 0$). So, if everyone has the same tastes, then the tax rates must balance the efficiency costs and the weighted compensation benefits.

To get more insight in the optimal tax structure, and to compare it with the existing literature, we focus next on two special cases. Section 4 starts with the simplest case possible: only income and no non-income factors. Afterwards, we add an exogenous tag to income in section 5.

4 Income only

We start with the simplest case possible: only income matters. The case is similar to Sheshinski (1972), but recall that agents differ in both abilities and tastes here. The system of first-order conditions of proposition 1 reduces to

$$(1-\alpha_0)\alpha_0\delta_0\frac{t_0}{1-t_0} = r\{(1-\alpha_0)^2(1-t_0)\,\sigma_{00}^{\theta} - (\alpha_0\delta_0)^2\,t_0\sigma_{00}^{\gamma}\}. \tag{1}$$

Responsibility for tastes implies that taxation has a responsibility cost. The optimal income tax rate will therefore be smaller in the presence of taste heterogeneity compared to models with heterogeneity in abilities only. Still, some of the classical comparative statics remain unchanged; see, e.g., Piketty and Saez (2013). In particular, the optimal income tax rate lies between 0 and 1, decreases with the elasticity of effort δ_0, and increases with inequality aversion r.

Fairness requires a higher sensitivity to ability differences compared to taste differences. The source of heterogeneity therefore plays a role for the optimal income tax rate. More ability heterogeneity σ_{00}^{θ} leads to higher income taxes, while more taste heterogeneity σ_{00}^{γ} implies lower taxes. The latter effect also occurs in Lockwood and Weinzierl's (2012) most plausible specification. Both results can also be combined to state that the optimal income tax rate must increase with the signal-to-noise ratio $\sigma_{00}^{\theta}/\sigma_{00}^{\gamma}$ (in case numerator and denominator change in opposite ways). The signal-to-noise ratio plays a similar role in the political economy model of Alesina and Angeletos (2005) and in the optimal tax model of Su and Judd (2006).

More control—a higher α_0—implies a lower income tax rate. To the best of our knowledge, this result is new in optimal tax models. It mirrors the political economy equilibria of Piketty (1995), Alesina and Angeletos (2005), and Bénabou and Tirole (2006) where a higher belief in control coincides with a lower tax rate.

Proposition 2 summarises the different theoretical results.

Proposition 2. The optimal tax rate t_0^* on income

1. lies in between the extremes of no taxation and complete taxation, i.e., $0 < t_0^* < 1$;

2. decreases with the elasticity of effort δ_0, from complete taxation in the case of perfectly inelastic effort ($t_0^* \to 1$ if $\delta_0 \to 0$) to no taxation in the case of perfectly elastic effort ($t_0^* \to 0$ if $\delta_0 \to +\infty$);

3. increases with the inequality aversion r, from no taxation if the planner is inequality neutral ($t_0^* \to 0$ if $r \to 0$) to partial taxation if the planner only cares about inequality ($0 < t_0^* < 1$ if $r \to +\infty$);

4. increases with ability heterogeneity σ_{00}^{θ}, from no taxation if everyone has the same ability ($t_0^* \to 0$ if $\sigma_{00}^{\theta} \to 0$) to complete taxation if ability becomes extremely heterogeneous ($t_0^* \to 1$ if $\sigma_{00}^{\theta} \to +\infty$);

5. decreases with taste heterogeneity σ_{00}^{γ}, from partial taxation if everyone has the same taste ($0 < t_0^* < 1$ if $\sigma_{00}^{\gamma} \to 0$) to zero taxation if taste becomes extremely heterogeneous ($t_0^* \to 0$ if $\sigma_{00}^{\gamma} \to +\infty$);

6. decreases with the degree of control α_0, from complete taxation if income cannot be controlled ($t_0^* \to 1$ if $\alpha_0 \to 0$) to no taxation if income is fully controlled ($t_0^* \to 0$ if $\alpha_0 \to 1$).

5 Adding a tag

Suppose that, in addition to income, there is also a tag, an observable non-controllable non-income factor, as in Akerlof (1978). The tag influences well-being directly as a non-income factor, but, in addition, it may also correlate with—and thus signal—unobserved earnings ability. Taxing or subsidising the tag has therefore two potential effects. It may reduce well-being differences that are directly caused by the tag (via β_1). In addition, it may also help reduce differences in well-being caused by differences in earnings ability because the tag is correlated with earnings ability (via σ_{10}^{θ}). The latter is called tagging. It has no efficiency cost (the tag is non-controllable), but it is imperfect (the tag is not a perfect signal of earnings ability).

In case of income and a single tag, the system of first-order conditions reduces to

$$\frac{\alpha_0\delta_0\zeta t_0}{1-t_0} = r\{(1-\alpha_0)\left((1-t_0)(1-\alpha_0)\sigma_{00}^{\theta} + (\beta_1 - t_1)\sigma_{10}^{\theta}\right) - (\alpha_0\delta_0)^2 t_0 \sigma_{00}^{\gamma}\}, \quad (2)$$

$$0 = (1-t_0)(1-\alpha_0)\sigma_{01}^{\theta} + (\beta_1 - t_1)\sigma_{11}^{\theta}, \quad (3)$$

with $\zeta = (1-\alpha_0) + \beta_1$ here.

5.1 The tax on income in the presence of a tag

In the previous section, the optimal income tax rate t_0^* turned out to be smaller in the presence of taste heterogeneity. We explain why it will be even smaller in the presence of a tag. If there is neither a direct effect of the tag on well-being ($\beta_1 \to 0$) nor an indirect signalling effect ($\sigma^\theta_{01} \to 0$), then equation (3) tells us that taxing or subsidising the tag makes no sense (i.e., $t_1 \to 0$ is optimal). In this case, equation (2) reduces to equation (1) and the optimal income tax rates must coincide. In addition, the comparative statics tell us that the optimal income tax rate (in the presence of a tag) decreases with the direct effect of the tag on well-being (β_1) and with the absolute value of the covariance between the tag and earnings ability ($|\sigma^\theta_{01}|$). Combining both results, the optimal tax rate on income will be lower in the presence of a tag.

Although the optimal tax rate on income t_0^* will be generically lower compared to the previous section, the comparative statics in proposition 2 remain valid. In addition, the correlation between the tag and unobserved earnings ability plays an interesting role. In the limiting cases of perfect correlation ($(\sigma^\theta_{01})^2 \to \sigma^\theta_{00}\sigma^\theta_{11}$), the tax rate on income t_0^* reduces to zero and all taxation can be done via t_1^*, the tax on the tag. This stands to reason because in these cases the tag is a perfect signal of unobserved earnings ability and, being non-controllable, it is a superior tax base as it can be taxed without efficiency cost. Finally, the optimal tax rate on income t_0^* increases with the variance of the tag σ^θ_{11}. If the tag becomes more noisy, tagging becomes less interesting relative to taxing income, and the income tax rate therefore increases.

Proposition 3 collects the different results for the optimal tax rate on income in the presence of a tag.

PROPOSITION 3. The optimal tax rate t_0^* on income in the presence of a tag

1. satisfies all properties of proposition 2;
2. decreases with the direct effect of the tag on well-being β_1;
3. follows an inverse U-shaped pattern with respect to the covariance between the tag and earnings ability σ^θ_{01}, starting and ending at no taxation in case of perfect correlation ($t_0^* \to 0$ if $(\sigma^\theta_{01})^2 \to \sigma^\theta_{00}\sigma^\theta_{11}$) and reaching a maximum in case of no correlation ($\sigma^\theta_{01} = 0$);
4. increases with the variance of the tag σ^θ_{11}.

5.2 The tax on the tag

The second first-order condition can be rewritten as

$$t_1 = \beta_1 + (1 - t_0)(1 - \alpha_0)\sigma^\theta_{01}/\sigma^\theta_{11}. \tag{4}$$

The right-hand side consists of two parts, one dealing with the direct effect of the tag and the other with the indirect signalling effect. In the absence of a signalling effect ($\sigma_{01}^{\theta} \to 0$), the optimal (read: fair) tax on the tag is equal to β_1, i.e., the direct effect of the tag should be fully taxed away. In the absence of a direct effect of the tag on well-being ($\beta_1 \to 0$), the optimal (read: efficient) tax on the tag reduces to $t_1 = (1 - t_0)(1 - \alpha_0)\sigma_{01}^{\theta}/\sigma_{11}^{\theta}$. Because $1 - t_0 > 0$ in the optimum (proposition 2, point 1 and proposition 3, point 1), the tax on the tag will be positive (negative) if the tag signals a higher (lower) unobserved ability to earn.

To discuss the comparative statics for t_1^*, the tax on the tag, we assume—without loss of generality—a positive correlation between the tag and unobserved earnings ability. A higher cost of income taxation δ_0, less heterogeneity in earnings ability σ_{00}^{θ}, or more heterogeneity in tastes for earnings effort σ_{00}^{γ} implies that using an income tax becomes relatively less interesting compared to tagging. The tax on the tag will therefore be higher in these cases. The higher the inequality aversion r, the lower the tax on the tag. Although counterintuitive at first sight, recall that the inequality aversion only affects the tax on the tag via the tax t_0 on income in equation (4). A higher inequality aversion leads to a higher income tax that in turn reduces the indirect effect of the tag on well-being via *net* income. The effect of control on tagging is not clear a priori. Proposition 2 (point 6) and proposition 3 (point 1) tell us only that in the extreme cases of no control and full control the term $(1 - t_0)(1 - \alpha_0)$ in equation (4) is equal to zero. The tax on the tag must then be equal to β_1. Finally, the tax on the tag will be higher, the higher the signalling value of the tag for unobserved ability σ_{01}^{θ} and the lower the noise of the tag, measured by its variance σ_{11}^{θ}. So, the tax on the tag will be higher the higher the signal-to-noise ratio $\sigma_{01}^{\theta}/\sigma_{11}^{\theta}$ (in case numerator and denominator move in opposite ways).

Proposition 4 summarises the comparative statics for t_1.

PROPOSITION 4. The optimal tax rate t_1^* on the tag

1. will be larger (resp. smaller) than β_1, if the covariance σ_{01}^{θ} is positive (resp. negative);

2. increases (resp. decreases) with the income elasticity δ_0 if the covariance σ_{01}^{θ} is positive (resp. negative);

3. decreases (resp. increases) with ability heterogeneity for earnings σ_{00}^{θ} if the covariance σ_{01}^{θ} is positive (resp. negative);

4. increases (resp. decreases) with taste heterogeneity for earnings σ_{00}^{γ} if the covariance σ_{01}^{θ} is positive (resp. negative);

5. decreases (resp. increases) with the inequality aversion r if the covariance σ_{01}^{θ} is positive (resp. negative);

6. is equal to β_1 if there is no control over income and if there is full control over income ($t_1^* \to \beta_1$, if either $\alpha_0 \to 0$ or $\alpha_0 \to 1$); the change of the tax rate with control is undefined in general;[13]

7. increases with the covariance σ_{01}^{θ};

8. decreases (resp. increases) with σ_{11}^{θ} if the covariance σ_{01}^{θ} is positive (resp. negative).

6 Testable conditions

To set the stage, notice that equation (4) is not testable: we do not observe the degree of control α_0 and the covariance between the tag and unobserved earnings ability σ_{01}^{θ}. The covariance between the tag $x_1 = \theta_1$ and gross income y is observable however, and, using equation (17) of the appendix, it can be written as

$$cov(x_1, y) = (1 - \alpha_0)\, \sigma_{01}^{\theta}. \tag{5}$$

Equation (4) then becomes

$$t_1 = \beta_1 + (1 - t_0)\, cov(x_1, y)/cov(x_1, x_1). \tag{6}$$

All terms are observable in principle. Suppose for example that the tag is physical ability. The tax rates t_0 and t_1 reflect the tax rate on earnings and invalidity/health benefits. The term β_1 is the willingness to pay for physical ability—the marginal rate of substitution between net income and physical ability. It can be estimated if one is willing to use happiness data as a proxy for well-being; see, e.g., Fleurbaey and Schokkaert (2012).[14] The covariance $cov(x_1, y)$ between physical ability and gross income and the variance of physical ability $cov(x_1, x_1)$ can be easily estimated. In particular, the ratio

[13] Simulations suggest that t_1^* typically follows an inverse U-shaped (resp. U-shaped) pattern with respect to the degree of control if the covariance is positive (resp. negative).

[14] It has been argued that using self-reported happiness or subjective well-being data gives the researcher direct information on individual well-being and it is not necessary to rely on revealed preferences or estimated individual utilities; see, e.g., the survey by Di Tella and MacCulloch (2006).

$cov(x_1, y)/cov(x_1, x_1)$ is the ordinary least squares (OLS) estimate of the slope when regressing income y on physical ability x_1. Equation (6) tells us that the tax rate on the tag should be equal to the direct effect of physical ability on well-being augmented by the indirect effect of physical ability on well-being via its expected effect on net income.

In the remainder of this section, we generalise equation (6) to allow for several non-income factors. Afterwards, we discuss the empirical implementation and link it to the equality of opportunity literature.

6.1 The general case

Consider income and several non-income factors. We partition the set of non-income factors $\{1, 2, \ldots, J\}$ in the set of non-controllable factors $N = \{j|\alpha_j \to 0\}$ (assumed to be non-empty) and the set of partially controllable factors $P = \{j|\alpha_j > 0\}$. The first-order conditions for the non-controllable factors in proposition 1 are equal to[15]

$$\textstyle\sum_{k \in N} cov(x_j, x_k)(t_k - \beta_k) = (1 - t_0)cov(x_j, y) + \sum_{k \in P}(\beta_k - t_k)cov(x_j, x_k),\ j \in N. \quad (7)$$

Suppose we have data for n individuals in a country on gross incomes and non-income factors, collected in a $n \times 1$ vector y, a $n \times |N|$ matrix X_N for the non-controllable factors, and a $n \times |P|$ matrix X_P for the partially controllable factors. All data are assumed to be normalised (the mean is equal to zero). We can replace the population covariances in (7) by their sample equivalents to obtain (in matrix notation)

$$(X'_N X_N)(t_N - \beta_N) = (1 - t_0)X'_N y + X'_N X_P(\beta_P - t_P),$$

with $t' = (t'_N, t'_P)$ and $\beta' = (\beta'_N, \beta'_P)$ collecting the tax rates and the willingness to pay for the non-controllable and partially controllable factors. Because $X'_N X_N$ is invertible—the (non-controllable) factors are not perfectly correlated by assumption—we get

$$t_N = \beta_N + (1 - t_0)\underbrace{(X'_N X_N)^{-1} X'_N y}_{\beta_{N0}} + \underbrace{(X'_N X_N)^{-1} X'_N X_P}_{\beta_{NP}}(\beta_P - t_P). \quad (8)$$

The iconic term '$(X'_N X_N)^{-1} X'_N Z$' is the linear projection of the non-controllable factors on either gross income (if $Z = y$) or on the controllable factors (if $Z = X_P$). They

[15] Similar to equation (5), we use the fact that

$$cov(x_j, y) = (1 - \alpha_0)\, \sigma^{\theta}_{0j},$$

for all $j \in N$ and similarly

$$cov(x_j, x_k) = (1 - \alpha_k)\, \sigma^{\theta}_{kj},$$

for all $j \in N$ and $k \in P$.

can be estimated in a multiple linear regression of income and the partially controllable factors on the non-controllable factors, say

$$y = X_N \beta_{N0} + \epsilon_y, \tag{9}$$

$$X_P = X_N \beta_{NP} + \epsilon_P. \tag{10}$$

β_{N0} is the $|N| \times 1$ slope vector with typical element β_{j0} capturing the effect of the non-controllable factor $j \in N$ on gross income; β_{NP} is the $|N| \times |P|$ slope matrix with typical element β_{jk} capturing the effect of the non-controllable factor $j \in N$ on the controllable factor $k \in P$; and ϵ_y and ϵ_P are a vector and matrix of i.i.d. error terms.

Equation (8) tells us that the tax rate on each non-controllable factor $j \in N$ (the left-hand side) should be equal to the *total* expected effect of the non-controllable factor on well-being (the right-hand side). This total effect can be split up in the *direct* effect of each factor on well-being (captured by β_N) augmented by the sum of the *indirect* expected effects of each factor on well-being. This indirect expected effect can run via (1) net income (i.e., the expected effect of a non-controllable factor on gross income, captured by β_{N0}, multiplied by $1 - t_0$ to obtain the net effect on well-being) and via (2) the partially controllable non-income factors (i.e., the expected effect of a non-controllable factor on a partially controllable factor, captured by β_{NP}, multiplied by $\beta_P - t_P$ to obtain the net effect on well-being).

6.2 Implementation

All variables in the equation system (8) are observable and therefore the system is testable in principle. Rather than obtaining separate estimates of the different parameters, there is an easy direct way to put the theory to the test if one is willing to use happiness data as a proxy for utility. In the appendix we derive the indirect utility function in equation (19). It is (in vector notation) equal to

$$v = \texttt{constant} + c + X_N \beta_N + X_P \beta_P,$$

with v the $n \times 1$ vector of indirect utilities, c the $n \times 1$ vector of net incomes, and `constant` a vector containing the same constant for each individual.

Suppose we have happiness data collected in a $n \times 1$ vector h as a proxy for utility. We could specify a happiness regression

$$h = \texttt{constant} + \kappa c + X_N \kappa \beta_N + X_P \kappa \beta_P + \epsilon_h,$$

with ϵ_h a vector of error terms. We include a parameter κ to capture the effect of net income on happiness; therefore, we also multiplied the direct effects in $\beta = (\beta_N, \beta_P)$ with κ to keep their interpretation as the willingness to pay for the different non-income factors. Let T be a vector containing the same demogrant for each individual. Net income is in matrix notation equal to

$$c = (1 - t_0)y - T - X_N t_N - X_P t_P. \tag{11}$$

Replacing net income c by (11) in the happiness equation and—taking up T in the `constant`—the happiness equation becomes

$$h = \texttt{constant} + \kappa(1 - t_0)y + X_N\kappa(\beta_N - t_N) + X_P\kappa(\beta_P - t_P) + \epsilon_h. \tag{12}$$

Finally, using equations (9)-(10), we can rewrite the happiness equation as[16]

$$h = \texttt{constant} + X_N\kappa[\beta_N - t_N + (1 - t_0)\beta_{N0} + \beta_{NP}(\beta_P - t_P)] + \eta. \tag{13}$$

Recall equation (8). If the tax-benefit scheme is efficient and fair, then the term between squared brackets in equation (13), the total effect of non-controllable factors on well-being, should be equal to zero. This provides us with a simple test. First, regress happiness h on all non-controllable factors in X_N, i.e.,

$$h = \texttt{constant} + X_N\zeta_N + \eta,$$

with ζ_N the slopes for the non-controllable factors. Second, test the joint hypothesis that the slope vector ζ_N is equal to zero, and, in case it is rejected, test it separately for the different factors. Irrespective of the test results, the estimated slope vector ζ_N can provide valuable information about the total degree of compensation for different non-controllable factors in different countries and different time periods.

Three final comments are in order. First, the adopted responsibility cut at the beginning of our paper is to keep individuals responsible for their preferences (tastes). However, we end up here with regressions that belong to the rivalling control approach, in which individuals are responsible for factors under control; see, e.g., Ramos and Van de gaer (2012) and Roemer and Trannoy (2013) for overviews of the equality of opportunity literature. Two assumptions of the current model play a crucial role in reconciling

[16] The error term is η is defined as

$$\eta = \epsilon_h + \kappa(1 - t_0)\epsilon_y + \epsilon_P\kappa(\beta_P - t_P).$$

both approaches and are therefore worth repeating: (1) the independence assumption between unobserved abilities and tastes, and (2) the fact that optimal individual effort does not depend on ability. In this case, the two approaches coincide in our model.

Second, it is very likely that the tests will be rejected in many countries for many factors. One explanation could be that societies are only willing to reduce the indirect effect of some non-controllable factors on well-being, but not the direct effect, say, the suffering caused by the factor.[17] Or, societies may not fully grasp the complex correlation structure of the different factors and disregard therefore some of the indirect effects when designing tax schemes. To further investigate such possibilities, one could

1. combine equations (9) and (11) to obtain the regression

$$c = \texttt{constant} + X_N[(1 - t_0)\beta_{N0} - t_N] - X_P t_P + \epsilon. \tag{14}$$

   Regressing net income c on X_N and X_P allows therefore to test whether the *indirect effects* of the non-controllable factors *via income* are compensated.

2. combine equations (10) and (11) to obtain

$$c = \texttt{constant} + X_N[-t_N - \beta_{NP} t_P] + (1 - t_0)y + \epsilon, \tag{15}$$

   and thus regressing net income c on X_N and y allows to test whether the *indirect effects* of the non-controllable factors *via the non-income factors* are fully compensated.

3. combine equations (9), (10), and (11) to obtain

$$c = \texttt{constant} + X_N[(1 - t_0)\beta_{N0} - t_N - \beta_{NP} t_P] + \epsilon. \tag{16}$$

   Regressing net income c on X_N allows to test whether *all indirect effects* of the non-controllable factors are fully compensated.

4. regress happiness h on y, X_N, and X_P—as in equation (12)—to test whether $t_N = \beta_N$, i.e., whether *only the direct effect* of the tag is compensated.

This framework is related to the empirical approach to inequality of opportunity. In this literature, practitioners distinguish between exogenous circumstances (X_N in our notation) and endogenous effort (X_P in our notation). Bourguignon et al. (2007) for

[17] It could simply be too costly or even impossible to fully compensate for severe disabilities or pains.

example, consider father's and mother's education, father's occupation, race, and region of birth to be circumstances, while they assign (own) schooling attainment, migration, and labour market status as effort variables. The much discussed correlation between circumstances and effort in the empirical literature corresponds with the correlation between non-controllable and controllable factors in our setting. In addition, the decomposition into direct and indirect effects resembles Bourguignon et al. (2007)'s empirical decomposition.

Third, estimating inequality of opportunity—the relative share of total inequality that can be attributed to circumstances—typically starts from estimating equations (9), (14), or (16) depending on the specific application. The empirical results for most countries suggest that roughly a quarter to a third of income inequality can be attributed to non-controllable factors; see, e.g., the results surveyed by Ramos and Van de gaer (2012) and Roemer and Trannoy (2013). It has been acknowledged that these estimates provide only lower bounds as researchers do not observe true ability nor (tastes for) effort; see, e.g., Ferreira and Gignoux (2011) and Niehues and Peichl (2013). In addition, the resulting coefficients cannot be interpreted as causal; see, e.g., Roemer and Trannoy (2013, p. 81). A promising avenue for future research could combine the empirical framework suggested here with identification strategies that exploit siblings correlations; see, e.g., Björklund and Jäntti (2012).

7 Conclusion

Fairness plays a role in redistribution. Individuals want to compensate for misfortunes, but also allow each other to enjoy the fruits of their effort. Such fairness considerations have been introduced in political economy and optimal income tax models. We introduce fairness as a device to select among efficient tax-benefit schemes based on income and non-income factors under partial control.

In general, optimal tax rates weigh the marginal efficiency cost of taxation caused by tax distortions against the marginal net fairness benefit of taxation. The latter combines two effects. The marginal compensation benefit of taxation captures the fairness benefit that higher taxes reduce outcome differences between individuals with the same tastes, but different abilities. The marginal responsibility cost of taxation measures the fairness cost that higher taxes also reduce the outcome differences between individuals with the same abilities, but different tastes.

We also study two special cases in detail. In case only income is included in the model, we show, among other things, that the optimal income tax negatively depends on the degree of control over income and on the heterogeneity in abilities and tastes in opposite ways. If we also add a tag to income, then the same conditions hold for the optimal income tax while the optimal tax on the tag depends on its direct effect on well-being and on its correlation with the ability to earn income.

The theoretical analysis suggests lower taxes on income and higher taxes on non-controllable non-income factors. While taxes on gender, age, and race are forbidden *de jure* by anti-discrimination laws, many tax-benefit schemes contain such taxes *de facto*. For instance, most existing tax systems have at least some elements of (or even complete) joint taxation. This punishes the secondary earner—usually the wife—with higher marginal tax rates. Thus, it would be interesting to bring our theory to the data and to investigate how existing tax systems (explicitly or implicitly) tax the various factors.

While we also derive testable conditions for the general case and discuss the empirical implementation, we leave the empirical estimation for future research. Several problems arise, and any estimation will have to satisfactorily deal with these issues. Panel data and fixed effects are preferably used to capture unobserved heterogeneity. Fixed effects estimation requires the non-controllable factors to vary over time. While this is probably true for disability, it is not true for gender or parental background. Also factors like age—if included—may be problematic as it is notoriously difficult to disentangle cohort, time, and age effects; see, e.g., Deaton and Paxson (1994). It is a priori not clear, however, whether age should be included at all in the analysis, e.g., if the goal is to look at average life-cycle utility (see, e.g., Weinzierl, 2011). One must also be sure that the included factors are non-controllable, which is not always easy to say, think, e.g., of being an immigrant in a country. Even if these problems were solved, potential identification problems remain. Bias caused by omitted variables, for example, is likely, as it is not possible to observe true abilities and tastes. To tackle these issues, exploiting sibling correlations might be a fruitful avenue.

References

[1] Aaberge, R., and Colombino, U., 2012, Accounting for family background when designing optimal income taxes: a microeconometric simulation analysis, *Journal of Population Economics* 25(2), 741-761.

[2] Akerlof, G., 1978, The economics of “tagging” as applied to the optimal income tax, welfare programs, and manpower planning, *American Economic Review* 68(1), 8-19.

[3] Alesina, A., and Angeletos, G.-M., 2005, Fairness and redistribution, *American Economic Review* 95(4), 960-980.

[4] Alesina, A., Angeletos, G.-M., and Cozzi, G., 2013, Fairness and redistribution: reply, *American Economic Review* 103(1), 554-561.

[5] Alesina, A., Cozzi, G., and Mantovan, N., 2012, The evolution of ideology, fairness and redistribution, *The Economic Journal* 122, 1244-1261.

[6] Alesina, A., and Giuliano, P., 2010, Preferences for redistribution, in, Benhabib, J., Jackson, M., and Bisin, A., eds, *Handbook of Social Economics*, 93-131, North-Holland: Elsevier.

[7] Alesina, A., Glaeser, E., and Sacerdote, B., 2001, Why doesn't the United States have a European-style welfare state? *Brookings Papers on Economic Activity* 2, 187-277.

[8] Alesina, A., Ichino, A., and Karabarbounis, L., 2011, Gender based taxation and the division of family chores, *American Economic Journal: Economic Policy*, 3(2), 1-40.

[9] Bargain, O., Orsini, K., and Peichl, A., 2014, Comparing labor supply elasticities in Europe and the US: new results", forthcoming in *Journal of Human Resources.*

[10] Bastani, S., Blomquist, S., and Micheletto, L., 2013, The welfare gains of age related optimal income taxation, *International Economic Review* 54(4), 1219-1249.

[11] Bénabou, R., and Tirole, J., 2006, Belief in a just world and redistributive politics, *The Quarterly Journal of Economics* 121(2), 699-746.

[12] Björklund, A., and Jäntti, M., 2012, How important is family background for labor-economic outcomes?, *Labour Economics* 19(4), 465-474.

[13] Blomquist, S., and Micheletto, L., 2008, Age-related optimal income taxation, *Scandinavian Journal of Economics* 110(1), 45-71.

[14] Boadway, R., Marchand, M., Pestieau, P., and Racionero, M., 2002, Optimal redistribution with heterogeneous preferences for leisure, *Journal of Public Economic Theory* 4(4), 475-498.

[15] Boadway, R., 2012, *From Optimal Tax Theory to Tax Policy: Retrospective and Prospective Views*, Cambridge: MIT Press.

[16] Boskin, M., and Sheshinski, E., 1983, Optimal tax treatment of the family: married couples, *Journal of Public Economics* 20(3), 281-297.

[17] Bourguignon, F., Ferreira, F., and Menéndez, M., 2007, Inequality of opportunity in Brazil, *Review of Income and Wealth* 53(4), 585-618.

[18] Brett, C., 2007, Optimal nonlinear taxes for families, *International Tax and Public Finance* 14(3), 225-261.

[19] Choné, P., and Laroque, G., 2010, Negative marginal tax rates and heterogeneity, *American Economic Review* 100, 2532-2547.

[20] Cremer, H., Dellis, A. and Pestieau, P., 2003, Family size and optimal income taxation, *Journal of Population Economics*, 16(1), 37-54.

[21] Cremer, H., Gahvari, F., and Lozachmeur, J.-M., 2010, Tagging and income taxation: theory and an application, *American Economic Journal: Economic Policy* 2(1), 31-50.

[22] Cremer, H., Lozachmeur, J.-M., and Pestieau, P., 2012, Income taxation of couples and the tax unit choice, *Journal of Population Economics*, 25(2), 763-778.

[23] Deaton, A. and Paxson, C., 1994, Intertemporal choice and inequality, *Journal of Political Economy* 102(3), 437-467.

[24] Diamond, P., 1998, Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates, *American Economic Review* 88 (1), 83-95.

[25] Diamond, P., and Saez, E., 2011, The case for a progressive tax: from basic research to policy recommendations, *Journal of Economic Perspectives* 25(4), 165-190

[26] Di Tella, R., and Dubra, J., 2013, Fairness and redistribution: comment, *American Economic Review* 103(1), 549-553.

[27] Di Tella, R., and MacCulloch, R., 2006, Some uses of happiness data in economics, *Journal of Economic Perspectives* 20(1), 25-46.

[28] Fehr, E., and Schmidt, K.M., 2006, The economics of fairness, reciprocity and altruism: experimental evidence, in, Kolm, S., and Ythier, J.M., eds, *Handbook of the Economics of Giving, Reciprocity and Altruism*, volume 1, Foundations, North-Holland: Elsevier.

[29] Ferreira, G., and Gignoux, J., 2011, The measurement of inequality of opportunity: theory and an application to Latin America, *Review of Income and Wealth* 57(4), 622-657.

[30] Fleurbaey, M., and Maniquet, F., 2006, Fair income tax, *Review of Economic Studies* 73(1), 55-83.

[31] Fleurbaey, M., and Maniquet, F., 2007, Help the low skilled or let the hardworking thrive? A study of fairness in optimal income taxation, *Journal of Public Economic Theory* 9(3), 467-500.

[32] Fleurbaey, M., and Schokkaert, E., 2012, Equity in health and health care, in Pauly, M., Mcguire, T., and Barros, P. (eds.), *Handbook of Health Economics*, volume 2, New York, Elsevier.

[33] Gaertner, W., and Schokkaert, E., 2011, *Empirical Social Choice*, Cambridge: Cambridge University Press.

[34] Immervoll, H., Kleven, H., Kreiner, C., and Verdelin, N., 2011, Optimal tax and transfer programs for couples with extensive labor supply responses, *Journal of Public Economics* 96(11-12), 1485-1500.

[35] Immonen, R., Kanbur, R., Keen, M., and Tuomala, M., 1998, Tagging and taxing: the optimal use of categorical and income information in designing tax/transfer schemes, *Economica* 65(258), 179-192.

[36] Jacquet, L., and Van de gaer, D., 2011, A comparison of optimal tax policies when compensation or responsibility matter, *Journal of Public Economics* 95 (11-12), 1248-1262.

[37] Kaplow, L., 2008, Optimal policy with heterogeneous preferences, NBER working paper 14170.

[38] Kleven, H., Kreiner, C., and Saez, E., 2009, The optimal income taxation of couples, *Econometrica*, 77 (2), 537-560.

[39] Konow, J., 2003, Which is the fairest one of all? A positive analysis of justice theories, *Journal of Economic Literature* 41(4), 1188-1239.

[40] Lockwood, B., and Weinzierl, M., 2012, De gustibus non est taxandum: theory and evidence on preference heterogeneity and redistribution, NBER Working Paper 17784.

[41] Luttens, R.I., and Ooghe, E., 2007, Is it fair to 'make work pay'?, *Economica* 74(296), 599-626.

[42] Mankiw, N.G., Weinzierl, M., and Yagan, D., 2009, Optimal taxation in theory and practice, *Journal of Economic Perspectives* 23(4), 147-174.

[43] Mankiw, N.G., and Weinzierl, M., 2010, The optimal taxation of height: a case study of utilitarian income redistribution, *American Economic Journal: Economic Policy* 2(1), 155-176

[44] Mirrlees, J., 1971, An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38(114), 175-208.

[45] Mirrlees J., 1972, Population policy and the taxation of family size, *Journal of Public Economics* 1(2), 169-198.

[46] Mueller, D., 2003, *Public Choice III*, Cambridge: Cambridge University Press.

[47] Niehues, J. and Peichl, A., 2014, Lower and upper bounds of unfair inequality: theory and evidence for Germany and the US, forthcoming in *Social Choice and Welfare*.

[48] Ooghe, E., and Peichl, A, 2010, Fair and efficient taxation under partial control: theory and evidence, IZA Discussion Paper 5388.

[49] Pigou, A.C., 1920, *The Economics of Welfare*, London: Macmillan and Co.

[50] Piketty, T., 1995, Social mobility and redistributive politics, *The Quarterly Journal of Economics* 110(3), 551-584.

[51] Piketty, T., and Saez, E., 2013, Optimal labor income taxation, in Auerbach, A., Chetty, R., Feldstein, M., and Saez, E. (eds.), *Handbook of Public Economics*, Volume 5, North-Holland: Elsevier.

[52] Ramos, X., and Van de gaer, D., 2012, Empirical approaches to inequality of opportunity: principles, measures, and evidence, IZA Discussion Paper 6672.

[53] Roemer, J., Aaberge, R., Colombino, U., Fritzell, J., Jenkins, S.P., Lefranc, A., Marx, I., Page, M., Pommer, E., Ruiz-Castillo, J., San Segundo, M.J., Tranaes, T., Trannoy, A., Wagner, G., Zubiri, I., 2003, To what extent do fiscal regimes equalize opportunities for income acquisition among citizens? *Journal of Public Economics* 87(3-4), 539-565.

[54] Roemer, J., and Trannoy, A, 2013, Equality of opportunity, Cowles Foundation Discussion Paper No. 1921, forthcoming in *Handbook of Income Distribution.*

[55] Salanié, B., 2002, Optimal demogrants with imperfect tagging, *Economics Letters* 75(3), 319-324.

[56] Salanié, B., 2003, *The Economics of Taxation*, MIT Press.

[57] Schokkaert, E., Van de gaer, D., Vandenbroucke, F., and Luttens, R., 2004, Responsibility sensitive egalitarianism and optimal linear income taxation, *Mathematical Social Sciences* 48(2), 151-182.

[58] Schroyen, F., 2003, Redistributive taxation and the household: the case of individual filings, *Journal of Public Economics*, 87, 2527-2547.

[59] Sheshinski, E., 1972, The optimal linear income-tax, *The Review of Economic Studies* 39(3), 297-302.

[60] Su, C.-L., and Judd, K.L., 2006, Optimal income taxation with multidimensional taxpayer types, *Computing in Economics and Finance* 471, Society for Computational Economics.

[61] Weinzierl, M., 2011, The surprising power of age-dependent taxes, *Review of Economic Studies* 78 (4), 1490-1518.

Principles underlying the social preference relation

We prove that the social preference relation satisfies the Pareto principle, compensation, and responsibility. We start with the individual utility maximisation problem, being

$$\max_e U(c,x,e) = c + \textstyle\sum_{j=1}^J \beta_j x_j - \sum_{j=0}^J \frac{\delta_j}{\exp\gamma_j}\exp(\frac{e_j}{\delta_j})$$

subject to the following constraints

$$\begin{aligned} c &\leq (1-t_0)y - T - \textstyle\sum_{j=1}^J t_j x_j, \\ y &= \alpha_0 e_0 + (1-\alpha_0)\theta_0, \\ x_j &= \alpha_j e_j + (1-\alpha_j)\theta_j \text{ for all } j = 1,2,\ldots,J. \end{aligned}$$

Define $\beta_0 = 1$. The $J+1$ first order conditions are

$$(\beta_j - t_j)\alpha_j - \frac{1}{\exp\gamma_j}\exp(\frac{e_j}{\delta_j}) = 0,$$

and lead to optimal effort choices[18]

$$e_j^* = \delta_j[\ln(\beta_j - t_j)\alpha_j + \gamma_j] \text{ for all } j = 0,1,\ldots,J.$$

The corresponding gross income, non-income factors, and consumption are equal to

$$\begin{aligned} y^* &= \alpha_0 e_0^* + (1-\alpha_0)\,\theta_0 \\ &= \alpha_0\delta_0[\ln((\beta_0 - t_0)\,\alpha_0) + \gamma_0] + (1-\alpha_0)\,\theta_0, && (17) \\ x_j^* &= \alpha_j e_j^* + (1-\alpha_j)\,\theta_j \\ &= \alpha_j\delta_j[\ln\left(\left(\beta_j - t_j\right)\alpha_j\right) + \gamma_j] + (1-\alpha_j)\,\theta_j, && (18) \\ c^* &= (\beta_0 - t_0)y^* - T - \textstyle\sum_{j=1}^J t_j x_j^*. \end{aligned}$$

From now on we use (T,t) rather than τ to refer to a tax scheme. Indirect utility is denoted $V(T,t,\theta,\gamma)$ and is equal to

$$\begin{aligned} V(T,t,\theta,\gamma) &= c^* + \textstyle\sum_{j=1}^J \beta_j x_j^* - \sum_{j=0}^J \frac{\delta_j}{\exp\gamma_j}\exp(\frac{e_j^*}{\delta_j}), \\ &= c^* + \textstyle\sum_{j=1}^J \beta_j x_j^* - \sum_{j=0}^J \delta_j(\beta_j - t_j)\alpha_j, && (19) \\ &= -T + (\beta_0 - t_0)y + \textstyle\sum_{j=1}^J(\beta_j - t_j)x_j^* - \sum_{j=0}^J \delta_j(\beta_j - t_j)\alpha_j, \\ &= \kappa(T,t) + \textstyle\sum_{j=0}^J(\beta_j - t_j)\alpha_j\delta_j\gamma_j + \sum_{j=0}^J(\beta_j - t_j)\,(1-\alpha_j)\,\theta_j, \end{aligned}$$

[18] Because $t_j < \beta_j$ must hold (for all $j = 0,1,\ldots,J$) in the social optimum, the optimal effort choices are well-defined.

with

$$\kappa(T,t) = -T + \textstyle\sum_{j=0}^{J}(\beta_j - t_j)\alpha_j\delta_j[\ln\left(\left(\beta_j - t_j\right)\alpha_j\right) - 1]. \tag{20}$$

Well-being $\widehat{v} = v(T,t,\theta,\gamma)$ follows from equating

$$V(T,t,\theta,\gamma) = V(R_0, 0, (\widehat{v}, \widehat{v}, \ldots, \widehat{v}), \gamma),$$

leading to

$$v(T,t,\theta,\gamma) = \frac{\kappa(T,t) - \kappa(R_0,0) - \sum_{j=0}^{J} t_j\alpha_j\delta_j\gamma_j + \sum_{j=0}^{J}(\beta_j - t_j)\left(1-\alpha_j\right)\theta_j}{\sum_{j=0}^{J}\beta_j\left(1-\alpha_j\right)}. \tag{21}$$

We are now ready to prove the properties of the social ranking. First of all, for each taste vector γ, well-being v is a strictly increasing (affine) transformation of V. We indeed have

$$v(T,t,\theta,\gamma) = a(\gamma) + b \times V(T,t,\theta,\gamma),$$

with

$$a(\gamma) = \frac{-\kappa(R_0,0) - \sum_{j=0}^{J}\beta_j\alpha_j\delta_j\gamma_j}{\sum_{j=0}^{J}\beta_j\left(1-\alpha_j\right)} \text{ and } b = \frac{1}{\sum_{j=0}^{J}\beta_j\left(1-\alpha_j\right)} > 0.$$

Because welfare is strictly increasing in well-being and well-being is strictly increasing in indirect utility, the social preference relation satisfies the Pareto principle, i.e., a higher utility for everyone (and strictly higher for at least one individual) implies a (strictly) higher social welfare.

Second, the social welfare weight of an individual with type (θ,γ) is equal to the derivative of welfare w.r.t. well-being multiplied by the derivative of well-being w.r.t. income. In the current setting, we obtain

$$\frac{\exp(-rv(T,t,\theta,\gamma))}{\int_\theta\int_\gamma \exp(-rv(T,t,\theta,\gamma))dF\left(\theta\right)dG\left(\gamma\right)} \times b,$$

with $-\partial v(T,t,\theta,\gamma)/\partial T = b$ the marginal well-being of income. The *relative* social welfare weight of two individuals is thus inversely related to their well-being level. If two individuals have the same tastes, then the one with the lower well-being level will get priority, i.e., a transfer from the better off to the worse off—if it were feasible—improves social welfare as required by the compensation principle. In the laisser-faire—defined as $(T,t) = (R_0,0)$—individuals with the same abilities have the same well-being. Indeed, we have

$$v(R_0,0,\theta,\gamma) = v(R_0,0,\theta,\gamma') = \frac{\sum_{j=0}^{J}\beta_j\left(1-\alpha_j\right)\theta_j}{\sum_{j=0}^{J}\beta_j\left(1-\alpha_j\right)},$$

for all $R_0, \theta, \gamma, \gamma'$. As a consequence, such individuals have the same social welfare weight. If all individuals have the same abilities, then their social welfare weight is the same in the laisser-faire. Any redistribution would be both inefficient (distortive) and inequality-increasing, thus the laisser-faire will be optimal as required by the responsibility principle.

Proof of proposition 1

The planner chooses a tax scheme τ to maximise

$$W(\tau) = \phi^{-1}[\textstyle\int_\theta \int_\gamma \phi(v(T, t, \theta, \gamma)) dF(\theta)\, dG(\gamma)]$$

subject to the budget constraint

$$\textstyle\int_\theta \int_\gamma \tau(y^*(\tau, \theta, \gamma), x^*(\tau, \theta, \gamma)) dF(\theta)\, dG(\gamma) \geq R_0.$$

We rewrite the budget constraint and the welfare function on the basis of the assumptions made. For ease of exposition, we define $\beta_0 = 1$.

With linear taxes, the budget constraint is equal to

$$T + \textstyle\int_\theta \int_\gamma (t_0 y^* + \sum_{j=1}^J t_j x_j^*) dF(\theta)\, dG(\gamma) \geq R_0,$$

with y^*, $x_j^*, j = 1, 2, \ldots, J$ defined in equations (17)-(18) and F and G multivariate normal distributions, with μ^θ and μ^γ the vector of means. Plugging in the different expressions, we can rewrite the budget constraint as

$$T + \textstyle\sum_{j=0}^J t_j \alpha_j \delta_j \ln\left(\left(\beta_j - t_j\right)\alpha_j\right) + \sum_{j=0}^J t_j \left(1 - \alpha_j\right)\mu_j^\theta + \sum_{j=0}^J t_j \alpha_j \delta_j \mu_j^\gamma \geq R_0.$$

Because efficiency requires the budget constraint to be satisfied with equality, the lump-sum tax T can be written as a function of the tax rates, i.e.,

$$T = R_0 - \textstyle\sum_{j=0}^J t_j \alpha_j \delta_j \ln\left(\left(\beta_j - t_j\right)\alpha_j\right) - \sum_{j=0}^J t_j \left(1 - \alpha_j\right)\mu_j^\theta - \sum_{j=0}^J t_j \alpha_j \delta_j \mu_j^\gamma. \quad (22)$$

Welfare is equal to

$$-\frac{1}{r} \ln[\textstyle\int_\theta \int_\gamma \exp(-rv(T, t, \theta, \gamma)) dF(\theta)\, dG(\gamma)],$$

with v defined in equation (21) and F and G multivariate normal distributions. Plugging in equation (21), welfare can be decomposed as $W = A + B + C$, with

$$\begin{aligned}
A &= \frac{\kappa(T,t) - \kappa(R_0,0)}{\zeta}, \\
B &= -\frac{1}{r}\ln[\textstyle\int_\theta \exp(\frac{-r\sum_{j=0}^{J}(\beta_j - t_j)(1-\alpha_j)\theta_j}{\zeta})dF(\theta)], \\
C &= -\frac{1}{r}\ln[\textstyle\int_\gamma \exp(\frac{r\sum_{j=0}^{J} t_j\alpha_j\delta_j\gamma_j}{\zeta})dG(\gamma)],
\end{aligned}$$

with $\zeta = \sum_{j=0}^{J}\beta_j(1-\alpha_j) > 0$ and κ defined in equation (20). We rewrite the different components A, B, and C.

Using equations (20) and (22), we directly get

$$A = \frac{\sum_{j=0}^{J} t_j\alpha_j\delta_j(1+\mu_j^\gamma) + \sum_{j=0}^{J} t_j(1-\alpha_j)\mu_j^\theta + \sum_{j=0}^{J}\beta_j\alpha_j\delta_j \ln\frac{\beta_j - t_j}{\beta_j}}{\zeta}.$$

To rewrite B and C, note that the moment-generating function of a normally distributed $x \sim N[\mu, \Sigma]$ is equal to

$$\int \exp(\textstyle\sum_{j=0}^{J} a_j x_j)dH(x) = \exp[\sum_{j=0}^{J} a_j\mu_j + \frac{1}{2}\sum_{i=0}^{J}\sum_{j=0}^{J} a_i a_j\sigma_{ij}].$$

We get

$$\begin{aligned}
B &= \textstyle\sum_{j=0}^{J}\frac{(\beta_j - t_j)(1-\alpha_j)\mu_j^\theta}{\zeta} - \frac{r\sum_{i=0}^{J}\sum_{j=0}^{J}(\beta_i - t_i)(1-\alpha_i)(\beta_j - t_j)(1-\alpha_j)\sigma_{ij}^\theta}{2\zeta^2}, \\
C &= -\textstyle\sum_{j=0}^{J}\frac{t_j\alpha_j\delta_j}{\zeta}\mu_j^\gamma - \frac{r\sum_{i=0}^{J}\sum_{j=0}^{J} t_i\alpha_i\delta_i t_j\alpha_j\delta_j\sigma_{ij}^\gamma}{2\zeta^2}.
\end{aligned}$$

Welfare $W = A+B+C$ is a function of tax rates $t = (t_0, t_1, \ldots, t_J)$ only. Maximising welfare leads to a system of first-order conditions of the form $\frac{\partial W}{\partial t_j} = \frac{\partial A}{\partial t_j} + \frac{\partial B}{\partial t_j} + \frac{\partial C}{\partial t_j} = 0$, with[19]

$$\begin{aligned}
\frac{\partial A}{\partial t_j} &= \frac{-\frac{\alpha_j\beta_j\delta_j}{\beta_j - t_j} + \alpha_j\delta_j\left(1+\mu_j^\gamma\right) + (1-\alpha_j)\mu_j^\theta}{\zeta}, \\
\frac{\partial B}{\partial t_j} &= -\frac{(1-\alpha_j)\mu_j^\theta}{\zeta} + \frac{r(1-\alpha_j)\sum_{k=0}^{J}(\beta_k - t_k)(1-\alpha_k)\sigma_{kj}^\theta}{\zeta^2}, \\
\frac{\partial C}{\partial t_j} &= -\frac{\alpha_j\delta_j\mu_j^\gamma}{\zeta} - \frac{r\alpha_j\delta_j\sum_{k=0}^{J} t_k\alpha_k\delta_k\sigma_{kj}^\gamma}{\zeta^2}.
\end{aligned}$$

Putting everything together we obtain

$$\alpha_j\delta_j\frac{t_j}{\beta_j - t_j} = r\{\frac{1-\alpha_j}{\zeta}\textstyle\sum_{k=0}^{J}(\beta_k - t_k)(1-\alpha_k)\sigma_{kj}^\theta - \frac{\alpha_j\delta_j}{\zeta}\sum_{k=0}^{J} t_k\alpha_k\delta_k\sigma_{kj}^\gamma\},$$

for each $j = 0, 1, \ldots, J$, as required.

[19] The double sums in B and C are of the generic form $\sum_{i=0}^{J}\sum_{j=0}^{J}\varphi_i(t_i)\varphi_j(t_j)\sigma_{ij}$, and its partial derivative with respect to t_j is equal to $2\frac{\partial\varphi_j(t_j)}{\partial t_j}\sum_{k=0}^{J}\varphi_k(t_k)\sigma_{kj}$.

Proof of proposition 2

If only income matters, then the first-order condition is

$$-(1-\alpha_0)\alpha_0\delta_0\frac{t_0}{1-t_0} - r(\alpha_0\delta_0)^2 t_0\sigma_{00}^{\gamma} + r(1-\alpha_0)^2(1-t_0)\sigma_{00}^{\theta} = 0. \tag{23}$$

The proof of proposition 2 turns out to be a special case of proposition 3 (point 1). We will come back to it in the next section.

Proof of proposition 3

Suppose there are two variables, income y and an exogenous tag x_1 (thus, $\alpha_1 \to 0$). The first-order conditions reduce to

$$\begin{aligned} -\alpha_0\delta_0\frac{\zeta t_0}{1-t_0} - r\left(\alpha_0\delta_0\right)^2 t_0\sigma_{00}^{\gamma} + r\left(1-\alpha_0\right)\Big(\left(1-t_0\right)\left(1-\alpha_0\right)\sigma_{00}^{\theta} + \left(\beta_1 - t_1\right)\sigma_{10}^{\theta}\Big) &= 0, \\ \left(1-t_0\right)\left(1-\alpha_0\right)\sigma_{01}^{\theta} + \left(\beta_1 - t_1\right)\sigma_{11}^{\theta} &= 0, \end{aligned}$$

with $\zeta = 1 - \alpha_0 + \beta_1 > 0$. The second first-order condition requires

$$t_1 = \beta_1 + \left(1-t_0\right)\left(1-\alpha_0\right)\frac{\sigma_{01}^{\theta}}{\sigma_{11}^{\theta}}, \tag{24}$$

which can be plugged in in the first condition, to get

$$-\underbrace{(1-\alpha_0+\beta_1)}_{\zeta>0}\alpha_0\delta_0\frac{t_0}{1-t_0} - r\left(\alpha_0\delta_0\right)^2 t_0\sigma_{00}^{\gamma} + r\left(1-\alpha_0\right)^2\left(1-t_0\right)\left(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}\right) = 0. \tag{25}$$

Note three things. First, equation (25) does not depend on t_1 and therefore completely describes the solution for income tax rate t_0. Second, the term $(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta})$ is strictly positive because the squared correlation $(\sigma_{01}^{\theta})^2/\sigma_{00}^{\theta}\sigma_{11}^{\theta}$ is assumed to be strictly smaller than 1. Third, if the tag has no direct effect ($\beta_1 = 0$) or indirect effect ($\sigma_{01}^{\theta} = 0$) on well-being, then equation (25) reduces to equation (23); therefore proposition 3 also proves proposition 2 as a special case.

POINT 1. The optimal tax rate t_0^* on income satisfies the properties mentioned in proposition 2 (we call them points 1.1-1.6 in the sequel).

POINT 1.1. The optimal tax rate t_0^* on income lies in between the extremes of no taxation and complete taxation, i.e., $0 < t_0^* < 1$.

If $t_0 \leq 0$ in the optimum, then the left-hand side of equation (25) is strictly positive and the first-order condition cannot be satisfied; so $t_1 > 0$ must hold at the optimum. If

t_0 approaches 1, then $\frac{t_0}{1-t_0}$ approaches $+\infty$, the left-hand side of equation (25) becomes strictly negative, and the first-order condition cannot be satisfied; so also $t_0 < 1$ must hold at the optimum.

POINT 1.2. The optimal tax rate t_0^* on income decreases with the elasticity δ_0, ranging from complete taxation in the case of perfectly inelastic effort ($t_0^* \to 1$ if $\delta_0 \to 0$) to no taxation in the case of perfectly elastic effort ($t_0^* \to 0$ if $\delta_0 \to +\infty$).

If $\delta_0 \to 0$, then the first-order condition reduces to

$$r\left(1-\alpha_0\right)^2\left(1-t_0\right)\left(\sigma_{00}^\theta-(\sigma_{01}^\theta)^2/\sigma_{11}^\theta\right)=0,$$

which is satisfied iff $t_0 \to 1$. If $\delta_0 \to +\infty$, then first divide both sides of equation (25) by $(\delta_0)^2 > 0$ and consider the limiting case $\delta_0 \to +\infty$ to get

$$-r\left(\alpha_0\right)^2 t_0 \sigma_{00}^\gamma=0,$$

which is satisfied iff $t_0 \to 0$. The comparative statics show that taxes decrease with δ_0, because

$$\frac{dt_0}{d\delta_0}=-\frac{\frac{\partial(25)}{\partial\delta_0}}{\frac{\partial(25)}{\partial t_0}}=-\frac{-\alpha_0\frac{\zeta t_0}{1-t_0}-2r\delta_0\left(\alpha_0\right)^2 t_0\sigma_{00}^\gamma}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2}-r\left(\alpha_0\delta_0\right)^2\sigma_{00}^\gamma-r\left(1-\alpha_0\right)^2\left(\sigma_{00}^\theta-(\sigma_{01}^\theta)^2/\sigma_{11}^\theta\right)},$$

is negative indeed, given $0 < t_0 < 1$ in the optimum.

POINT 1.3. The optimal tax rate t_0^* on income increases with the inequality aversion r, ranging from no taxation if the planner is inequality neutral ($t_0^* \to 0$ if $r \to 0$) to partial taxation if the planner only cares about inequality ($0 < t_0^* < 1$ if $r \to +\infty$).

If $r \to 0$, then equation (25) reduces to

$$-\alpha_0\delta_0\frac{\zeta t_0}{1-t_0}=0,$$

which implies $t_0 \to 0$. To investigate the case $r \to +\infty$, divide first both sides of (25) by $r > 0$, take the limit $r \to +\infty$, and solve for t_0 to obtain

$$t_0=\frac{\left(1-\alpha_0\right)^2\left(\sigma_{00}^\theta-(\sigma_{01}^\theta)^2/\sigma_{11}^\theta\right)}{\left(\alpha_0\delta_0\right)^2\sigma_{00}^\gamma+\left(1-\alpha_0\right)^2\left(\sigma_{00}^\theta-(\sigma_{01}^\theta)^2/\sigma_{11}^\theta\right)},$$

which will typically result in partial taxation. The comparative statics are given by

$$\frac{dt_0}{dr}=-\frac{\frac{\partial(25)}{\partial r}}{\frac{\partial(25)}{\partial t_0}}=-\frac{-\left(\alpha_0\delta_0\right)^2 t_0\sigma_{00}^\gamma+\left(1-\alpha_0\right)^2\left(1-t_0\right)\left(\sigma_{00}^\theta-(\sigma_{01}^\theta)^2/\sigma_{11}^\theta\right)}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2}-r\left(\alpha_0\delta_0\right)^2\sigma_{00}^\gamma-r\left(1-\alpha_0\right)^2\left(\sigma_{00}^\theta-(\sigma_{01}^\theta)^2/\sigma_{11}^\theta\right)}.$$

We can rewrite the numerator, using equation (25), to obtain

$$\frac{dt_0}{dr} = -\frac{\alpha_0\delta_0 \frac{\zeta t_0}{1-t_0}\frac{1}{r}}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} - r\left(\alpha_0\delta_0\right)^2\sigma_{00}^{\gamma} - r\left(1-\alpha_0\right)^2\left(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}\right)},$$

which is positive, indeed, given $0 < t_0 < 1$ in the optimum.

POINT 1.4. The optimal tax rate t_0^* on income increases with ability heterogeneity σ_{00}^{θ}, ranging from no taxation if everyone has the same ability ($t_0^* \to 0$ if $\sigma_{00}^{\theta} \to 0$) to complete taxation if ability becomes extremely heterogeneous ($t_0^* \to 1$ if $\sigma_{00}^{\theta} \to +\infty$).

If $\sigma_{00}^{\theta} \to 0$ (and thus also $\sigma_{01}^{\theta} \to 0$), then equation (25) reduces to

$$-(1-\alpha_0+\beta_1)\alpha_0\delta_0\frac{t_0}{1-t_0} - r\left(\alpha_0\delta_0\right)^2 t_0\sigma_{00}^{\gamma} = 0$$

which leads to $t_0 \to 0$. If $\sigma_{00}^{\theta} \to +\infty$, then divide first both sides of (25) by $\sigma_{00}^{\theta} > 0$, take the limit $\sigma_{00}^{\theta} \to +\infty$, and equation (25) reduces to

$$r\left(1-\alpha_0\right)^2\left(1-t_0\right) = 0,$$

which implies $t_0 \to 1$. The comparative statics for t_0 w.r.t. σ_{00}^{θ} are equal to

$$\frac{dt_0}{d\sigma_{00}^{\theta}} = -\frac{\frac{\partial(25)}{\partial\sigma_{00}^{\theta}}}{\frac{\partial(25)}{\partial t_0}} = -\frac{r\left(1-\alpha_0\right)^2\left(1-t_0\right)}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} - r\left(\alpha_0\delta_0\right)^2\sigma_{00}^{\gamma} - r\left(1-\alpha_0\right)^2\left(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}\right)},$$

which is positive, given $0 < t_0 < 1$ in the optimum.

POINT 1.5. The optimal tax rate t_0^* on income decreases with taste heterogeneity σ_{00}^{γ}, ranging from partial taxation if everyone has the same taste ($0 < t_0^* < 1$ if $\sigma_{00}^{\gamma} \to 0$) to zero taxation if taste becomes extremely heterogeneous ($t_0^* \to 0$ if $\sigma_{00}^{\gamma} \to +\infty$).

If $\sigma_{00}^{\gamma} \to 0$, then equation (25) reduces to

$$-\alpha_0\delta_0\frac{\zeta t_0}{1-t_0} + r\left(1-\alpha_0\right)^2\left(1-t_0\right)\left(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}\right) = 0,$$

which can lead to any tax rate in between 0 and 1. If $\sigma_{00}^{\gamma} \to +\infty$, then divide first both sides of (25) by $\sigma_{00}^{\gamma} > 0$, take the limit $\sigma_{00}^{\gamma} \to +\infty$, and equation (25) reduces to

$$-r\left(\alpha_0\delta_0\right)^2 t_0 = 0,$$

which implies $t_0 \to 0$. Comparative statics are given by

$$\frac{dt_0}{d\sigma_{00}^{\gamma}} = -\frac{\frac{\partial(25)}{\partial\sigma_{00}^{\gamma}}}{\frac{\partial(25)}{\partial t_0}} = -\frac{-r\left(\alpha_0\delta_0\right)^2 t_0}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} - r\left(\alpha_0\delta_0\right)^2\sigma_{00}^{\gamma} - r\left(1-\alpha_0\right)^2\left(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}\right)},$$

which is negative, as required.

POINT 1.6. The optimal tax rate t_0^* on income decreases with the degree of control α_0, ranging from complete taxation if income cannot be controlled ($t_0^* \to 1$ if $\alpha_0 \to 0$) to no taxation if income is fully controlled ($t_0^* \to 0$ if $\alpha_0 \to 1$).

If $\alpha_0 \to 0$, equation (25) reduces to

$$r\,(1-t_0)\,(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}) = 0,$$

which implies $t_0 \to 1$. If $\alpha_0 \to 1$, equation (25) reduces to

$$-\delta_0 \frac{\beta_1 t_0}{1-t_0} - r\,(\delta_0)^2\, t_0 \sigma_{00}^{\gamma} = 0,$$

which is satisfied iff $t_0 \to 0$. The comparative statics for t_0 w.r.t. α_0 are given by

$$\frac{dt_0}{d\alpha_0} = -\frac{\frac{\partial eq(25)}{\partial \alpha_0}}{\frac{\partial eq(25)}{\partial t_0}} = -\frac{(\alpha_0-\zeta)\delta_0 \frac{t_0}{1-t_0} - 2r(\alpha_0\delta_0)\delta_0 t_0 \sigma_{00}^{\gamma} - 2r\,(1-\alpha_0)\,(1-t_0)\,(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta})}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} - r\,(\alpha_0\delta_0)^2\,\sigma_{00}^{\gamma} - r\,(1-\alpha_0)^2\,(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta})}.$$

We can multiply the numerator and denominator by $\alpha_0\zeta = \alpha_0\,(1-\alpha_0+\beta_1) > 0$ and use equation (25) to replace $\alpha_0\delta_0\frac{\zeta t_0}{1-t_0}$ by $-r\,(\alpha_0\delta_0)^2\, t_0\sigma_{00}^{\gamma} + r\,(1-\alpha_0)^2\,(1-t_0)\,\sigma_{00}^{\theta}(1-(\rho_{01}^{\theta})^2)$ in the numerator, to obtain (after some manipulation)

$$\frac{dt_0}{d\alpha_0} = -\frac{(\alpha_0+\zeta)r\,(\alpha_0\delta_0)^2\, t_0\sigma_{00}^{\gamma} + (\zeta+\beta_1\alpha_0)r\,(1-\alpha_0)\,(1-t_0)\,(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta})}{\alpha_0\zeta[\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} + r\,(\alpha_0\delta_0)^2\,\sigma_{00}^{\gamma} + r\,(1-\alpha_0)^2\,(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta})]},$$

which is negative, given $0 < t_0 < 1$ in the optimum.

POINT 2. The optimal tax rate t_0^* on income decreases with the direct effect of the tag on well-being β_1.

The comparative statics for t_0 w.r.t. β_1 is given by

$$\frac{dt_0}{d\beta_1} = -\frac{\frac{\partial eq(25)}{\partial \beta_1}}{\frac{\partial eq(25)}{\partial t_0}} = -\frac{-\alpha_0\delta_0\frac{t_0}{1-t_0}}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} - r\,(\alpha_0\delta_0)^2\,\sigma_{00}^{\gamma} - r\,(1-\alpha_0)^2\,(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta})},$$

which is negative, given $0 < t_0 < 1$ in the optimum.

POINT 3. The optimal tax rate t_0^* on income follows an inverse U-shaped pattern with respect to the covariance between the tag and earnings ability σ_{01}^{θ}, starting and ending at no taxation in case of perfect correlation ($t_0^* \to 0$ if $(\sigma_{01}^{\theta})^2 \to \sigma_{00}^{\theta}\sigma_{11}^{\theta}$) and reaching a maximum in case of no correlation ($\sigma_{01}^{\theta} = 0$).

First, note that in the case of perfect correlation ($(\sigma_{01}^{\theta})^2 \to \sigma_{00}^{\theta}\sigma_{11}^{\theta}$) equation (25) reduces to

$$\alpha_0\delta_0\frac{\zeta t_0}{1-t_0} - r\left(\alpha_0\delta_0\right)^2 t_0\sigma_{00}^{\gamma} = 0,$$

which indeed implies $t_0 \to 0$. The comparative statics for t_0 w.r.t. σ_{01}^{θ} are given by

$$\frac{dt_0}{d\sigma_{01}^{\theta}} = -\frac{\frac{\partial eq(25)}{\partial\sigma_{01}^{\theta}}}{\frac{\partial eq(25)}{\partial t_0}} = -\frac{-r\left(1-\alpha_0\right)^2\left(1-t_0\right)2\sigma_{01}^{\theta}/\sigma_{11}^{\theta}}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} - r\left(\alpha_0\delta_0\right)^2\sigma_{00}^{\gamma} - r\left(1-\alpha_0\right)^2\left(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}\right)}.$$

The sign of $dt_0/d\sigma_{01}^{\theta}$ is indeed inversely related to the sign of σ_{01}^{θ}, leading to an inverse U-shaped pattern.

Point 4. The optimal tax rate t_0^* on income increases with σ_{11}^{θ}.

The comparative statics for t_0 w.r.t. σ_{11}^{θ} is given by

$$\frac{dt_0}{d\sigma_{11}^{\theta}} = -\frac{\frac{\partial eq(25)}{\partial\sigma_{11}^{\theta}}}{\frac{\partial eq(25)}{\partial t_0}} = -\frac{r\left(1-\alpha_0\right)^2\left(1-t_0\right)(\sigma_{01}^{\theta}/\sigma_{11}^{\theta})^2}{-\frac{\alpha_0\delta_0\zeta}{(1-t_0)^2} - r\left(\alpha_0\delta_0\right)^2\sigma_{00}^{\gamma} - r\left(1-\alpha_0\right)^2\left(\sigma_{00}^{\theta} - (\sigma_{01}^{\theta})^2/\sigma_{11}^{\theta}\right)},$$

which is positive, given $0 < t_0 < 1$ in the optimum.

Proof of proposition 4

Point 1. The optimal tax rate t_1^* on the tag will be larger (resp. smaller) than β_1, if the covariance σ_{01}^{θ} is positive (resp. negative)

Using proposition 3 (point 1.1 in the proof) we must have $0 < t_0 < 1$. Using equation (24), this implies indeed that $t_1 \gtreqless \beta_1$ holds if $\sigma_{01}^{\theta} \gtreqless 0$.

Point 2. The optimal tax rate t_1^* on the tag increases (resp. decreases) with the income elasticity δ_0 if the covariance σ_{01}^{θ} is positive (resp. negative).

The comparative statics of t_1 w.r.t. δ_0 are

$$\frac{dt_1}{d\delta_0} = \underbrace{\frac{\partial eq(24)}{\partial\delta_0}}_{=0} + \frac{\partial eq(24)}{\partial t_0}\frac{dt_0}{d\delta_0} = -\left(1-\alpha_0\right)\frac{\sigma_{01}^{\theta}}{\sigma_{11}^{\theta}}\frac{dt_0}{d\delta_0},$$

with $\frac{dt_0}{d\delta_0}$ negative at the optimum (see proposition 3, point 1.2 in the proof). The sign of $\frac{dt_1}{d\delta_0}$ corresponds therefore with the sign of σ_{01}^{θ}.

Point 3. The optimal tax rate t_1^* on the tag decreases (resp. increases) with ability heterogeneity for earnings σ_{00}^{θ} if the covariance σ_{01}^{θ} is positive (resp. negative).

The comparative statics of t_1 w.r.t. σ^θ_{00} are

$$\frac{dt_1}{d\sigma^\theta_{00}} = \underbrace{\frac{\partial eq(24)}{\partial \sigma^\theta_{00}}}_{=0} + \frac{\partial eq(24)}{\partial t_0}\frac{dt_0}{d\sigma^\theta_{00}} = -\left(1-\alpha_0\right)\frac{\sigma^\theta_{01}}{\sigma^\theta_{11}}\frac{dt_0}{d\sigma^\theta_{00}},$$

with $\frac{dt_0}{d\sigma^\theta_{00}}$ positive at the optimum (see proposition 3, point 1.4 in the proof). The sign of $\frac{dt_1}{d\delta_0}$ is thus inversely related to the sign of σ^θ_{01}.

Point 4. The optimal tax rate t_1^* on the tag increases (resp. decreases) with taste heterogeneity for earnings σ^γ_{00} if the covariance σ^θ_{01} is positive (resp. negative).

The comparative statics of t_1 w.r.t. σ^γ_{00} are

$$\frac{dt_1}{d\sigma^\gamma_{00}} = \underbrace{\frac{\partial eq(24)}{\partial \sigma^\gamma_{00}}}_{=0} + \frac{\partial eq(24)}{\partial t_0}\frac{dt_0}{d\sigma^\gamma_{00}} = -\left(1-\alpha_0\right)\frac{\sigma^\theta_{01}}{\sigma^\theta_{11}}\frac{dt_0}{d\sigma^\gamma_{00}},$$

with $\frac{dt_0}{d\sigma^\gamma_{00}}$ negative at the optimum (see proposition 3, point 1.5 in the proof). The sign of $\frac{dt_1}{d\sigma^\gamma_{00}}$ corresponds therefore with the sign of σ^θ_{01}.

Point 5. The optimal tax rate t_1^* on the tag decreases (resp. increases) with the inequality aversion r if the covariance σ^θ_{01} is positive (resp. negative).

The comparative statics of t_1 w.r.t. r are

$$\frac{dt_1}{dr} = \underbrace{\frac{\partial eq(24)}{\partial r}}_{=0} + \frac{\partial eq(24)}{\partial t_0}\frac{dt_0}{dr} = -\left(1-\alpha_0\right)\frac{\sigma^\theta_{01}}{\sigma^\theta_{11}}\frac{dt_0}{dr},$$

with $\frac{dt_0}{dr}$ positive at the optimum (see proposition 3, point 1.3 in the proof). The sign of $\frac{dt_1}{dr}$ is thus inversely related to the sign of σ^θ_{01}.

Point 6. The optimal tax rate t_1^* on the tag is equal to β_1 if there is no control over income and if there is full control over income ($t_1^* \to \beta_1$, if either $\alpha_0 \to 0$ or $\alpha_0 \to 1$); the change of the tax rate with control is undefined in general.

If $\alpha_0 \to 0$, then $t_0 \to 1$ and if $\alpha_0 \to 1$, then $t_0 \to 0$ (see proposition 3, point 1.6 in the proof). In both cases we have $(1-t_0)(1-\alpha_0) \to 0$ and equation (24) tells us that $t_1 \to \beta_1$. The comparative statics for t_1 w.r.t. α_0 are

$$\frac{dt_1}{d\alpha_0} = \frac{\partial eq(24)}{\partial \alpha_0} + \frac{\partial eq(24)}{\partial t_0}\frac{dt_0}{d\alpha_0} = -(1-t_0)\left(1+\frac{1-\alpha_0}{1-t_0}\frac{dt_0}{d\alpha_0}\right)\frac{\sigma^\theta_{01}}{\sigma^\theta_{11}},$$

with $\frac{dt_0}{d\alpha_0}$ negative (see proposition 3, point 1.6 in the proof). The sign of $\frac{dt_1}{d\alpha_0}$ is not defined in general.

POINT 7. The optimal tax rate t_1^* on the tag increases with σ_{01}^θ.

The comparative statics for t_1 w.r.t. σ_{01}^θ are

$$\frac{dt_1}{d\sigma_{01}^\theta} = \frac{\partial eq(24)}{\partial\sigma_{01}^\theta} + \frac{\partial eq(24)}{\partial t_0}\frac{dt_0}{d\sigma_{01}^\theta} = \frac{(1-t_0)(1-\alpha_0)}{\sigma_{11}^\theta} - (1-\alpha_0)\frac{\sigma_{01}^\theta}{\sigma_{11}^\theta}\frac{dt_0}{d\sigma_{01}^\theta},$$

with the sign of $\frac{dt_0}{d\sigma_{01}^\theta}$ being inversely related to the sign of σ_{01}^θ (see proposition 3, point 3), thus $\frac{\sigma_{01}^\theta dt_0}{d\sigma_{01}^\theta}$ must be negative. Using $0 < t_0 < 1$ at the optimum (see proposition 3, point 1.1 in the proof) we obtain that $\frac{dt_1}{d\sigma_{01}^\theta}$ is positive.

POINT 8. The optimal tax rate t_1^* on the tag decreases (resp. increases) with σ_{11}^θ if the covariance σ_{01}^θ is positive (resp. negative).

The comparative statics for t_1 w.r.t. σ_{11}^θ are

$$\frac{dt_1}{d\sigma_{11}^\theta} = \frac{\partial eq(24)}{\partial\sigma_{11}^\theta} + \frac{\partial eq(24)}{\partial t_0}\frac{dt_0}{d\sigma_{11}^\theta} = -\sigma_{01}^\theta[\frac{(1-t_0)(1-\alpha_0)}{(\sigma_{11}^\theta)^2} + \frac{1-\alpha_0}{\sigma_{11}^\theta}\frac{dt_0}{d\sigma_{11}^\theta}],$$

with the sign of $\frac{dt_0}{d\sigma_{11}^\theta}$ being positive (see proposition 3, point 4). The term between squared brackets is positive, and the sign of $\frac{dt_1}{d\sigma_{11}^\theta}$ is inversely related to the sign of σ_{01}^θ.