

Redistributive Taxation in a Roy Model*

Casey Rothschild
Wellesley College

Florian Scheuer
Stanford University

May 2012

Abstract

We consider optimal redistribution in a model where individuals can self-select into one of several possible sectors based on heterogeneity in a multidimensional skill vector. We show that when the government does not observe the sectoral choice or underlying skills of its citizens, the constrained Pareto frontier can be implemented with a single non-linear income tax. We derive formulas for this optimal tax schedule. If sectoral inputs are complements, a many-sector model with self-selection leads to optimal income taxes that are less progressive than the corresponding taxes in a standard single-sector model under natural conditions. However, they are more progressive than in a multi-sector model without occupational choice or without overlapping wage distributions, as in Stiglitz (1982).

*Email addresses: crothsch@wellesley.edu, scheuer@stanford.edu.

1 Introduction

Since Borjas (1987), the Roy (1951) model of self-selection has been a workhorse model in labor economics. Its essential characteristic is that individuals have a range of sectors to choose to work in and self-select into the one which affords them the greatest returns. For example, Borjas (2002) considers the implications of self-selection into the government versus non-government sectors. This sort of self-selection has obvious implications for the implementation of redistributive income taxation. It is surprising, then, that these implications have not been studied formally heretofore. This paper takes a step towards understanding these implications by studying optimal Mirrleesian income taxation in a two-sector Roy model.

Incorporating self-selection à la Roy in an optimal income tax framework à la Mirrlees raises some challenges. In particular, in the Mirrleesian approach, the government computing an optimal income tax effectively “screens” individuals based on their unobserved skill (or wage). When there are multiple sectors in which any given individual can choose to work, the underlying “skill” is multi-dimensional: any given individual has a skill for each possible sector. It is well known that multi-dimensional screening problems are challenging (Rochet and Choné, 1998). We show that the particular screening problem that is raised by considering optimal taxation in a many-sector model is tractable, however, despite the fact that the underlying skill is multi-dimensional.

We do this by showing that the most general direct mechanism for allocating consumption and effort to individuals in such a model is to use a single non-linear income tax. Since income is monotonic in wages, this allows us to use the single-dimensional screening tools developed by Mirrlees (1971), Diamond (1998) and others, with one important difference: the wage distribution will, in general, be *endogenous* when there is more than one sector to choose from. This is because the productivity of effort in any given sector will, in general, depend on the aggregate effort expended in that and in other sectors. (Agriculture will be more productive when there is a well-functioning transportation sector, and vice versa, for example.) We characterize optimal taxes accounting for this endogeneity and show that under quite natural and general assumptions it implies a force for *less* progressive taxation relative to a world with a single sector and an exogenous wage distribution.

The basic intuition for this result can be understood by considering the following simple framework. There are two sectors in the economy, a “blue collar” and “white collar” sector, and individuals are free to choose to work in either sector. The government is choosing an income tax system. It has a strict desire to redistribute from high income to

low income individuals, but it does not care about sectoral choice *per se*. Moreover, for administrative, informational, or political reasons, the income tax system chosen by the government cannot distinguish between the two sectors. Taxes are based only on *incomes earned*: a white collar and blue collar worker who earn the same income y will pay the same tax $T(y)$.

With a linear production technology (i.e., when white and blue collar work are perfect substitutes), the fact that there are two sectors would be irrelevant unless the government has an intrinsic preference for workers in one of the sectors. Individuals would choose to work in the sector in which they are more productive (as reflected by their wage). With linear technology, sectoral choices and the wage distribution would be independent of tax policy. Since the government does not care about sectoral choice *per se*, the optimal tax would therefore be exactly the same as it would be in a single-sector Mirrleesian framework with the same wage distribution.

Contrast this with the case in which the two sectors are gross complements. In this case, sectoral choices and wages are endogenous to tax policy. Raising taxes at income levels that are dominated by white collar workers, for example, will differentially discourage white collar effort. Under standard assumptions about aggregate productivity, this will raise the marginal productivity of white collar effort (by diminishing marginal products within a given sector) and lower the marginal productivity of blue collar effort (by complementarity across sectors). Such a tax therefore *indirectly redistributes* from blue to white collar workers.

Suppose now that the blue collar sector is the low-income sector, in the sense that the fraction of white collar workers is higher at higher income levels. Then this “indirect redistribution” channel will lead the government to choose a tax system which is less progressive than it would choose in a Mirrleesian world with exogenous wages: Lowering taxes on high earners will differentially spur effort in the white collar sector, which will indirectly redistribute from the disproportionately wealthy white collar workers to the disproportionately poor blue collar workers by raising blue collar wages and lowering white collar wages. Similarly, raising taxes on lower earners will differentially discourage effort in the blue collar sector, again indirectly redistributing from relatively rich white collar workers to relatively poor blue collar workers. The presence of this indirect redistribution channel therefore implies less progressive taxes than in a model with exogenous wages. For example, it implies an optimal top income tax rate that will generally be negative when the skill distribution is bounded—i.e., strictly below the well known zero top rate results.¹ Note that these conclusions would be further reinforced if the gov-

¹There is no intrinsic difference between the two sectors here except that one is assumed to be the low-

ernment also had a *direct* preference for redistribution towards blue collar workers, at any given income level: the indirect redistribution channel would push against any incentive this provides for increased progressivity.

It is important to note that this result does not say taxes should be regressive. Rather, it says that optimal taxes will be *less* progressive than they would be in the alternative allocation that would obtain if the endogeneity of wages implied by a multi-sector Roy model were neglected. Deriving this result requires that we further develop a formal notion for this alternative allocation. We use the notion of a self-confirming policy equilibrium (SCPE) for this (see Rothschild and Scheuer (2011) for a discussion of an SCPE in a different context). A SCPE here is a tax system (and allocations of effort and consumption) that would emerge in the same economy if the government *naively* believed that it was operating in a standard world with exogenous wages. In such a world, a government would, following Saez (2001), infer an underlying skill distribution from the income distribution it observes given an existing tax system. Taking this skill distribution as exogenous, it would then compute the optimal income tax system. In an SCPE, this newly computed optimal income tax system would coincide with the existing tax system, thus “confirming” its optimality. Our results show that taxes in such an SCPE are not, in fact, optimal in a multi-sector economy, since the wage-cum-skill distribution is *not*, in fact, exogenous. In particular, the optimal taxes would be less progressive.

Most closely related to our analysis is Stiglitz (1982), who considers optimal nonlinear taxation in a two-type model with endogenous wages, but without occupational choice. He also shows that the optimal top marginal tax rate is negative when the efforts of the two types are complements and redistribution is from the high to the low wage earners.² Our model differs in two significant ways. First, our continuous type model allows us to compare the progressivity of the entire tax schedule across different policy regimes, rather than just the top marginal tax rates. Second, we show that the general Roy model we consider gives rise to *extra* effects, which result from (i) endogenous occupational choice and (ii) the fact that our model with continuous types generates overlapping wage distributions in the two sectors, whereas a discrete type model generically – and somewhat unrealistically – rules out workers in different sectors earning the same wage.

We show that these extra effects mitigate the general equilibrium effects of taxation found in Stiglitz (1982) and therefore make optimal taxes *more* progressive than in a dis-

income and the other the high-income sector. In particular, the same results about the shape of the income tax schedule would obtain if the blue collar sector was the high-income rather than the low-income sector.

²See also Naito (1999), who focuses on the role of sector-specific taxes in such a two-type model. He points out that manipulating wages can relax incentive constraints and therefore be desirable even if it introduces production inefficiencies.

crete type model without occupational choice. To understand why, suppose we reduce taxes at the top to increase the effort of the top earners. This is desirable insofar as it indirectly redistributes from high to low incomes by raising the wages of workers in the low wage sector and lowering the wages of workers in the high wage sector. When there is endogenous occupational choice, however, there is an additional effect: the change in wages leads some individuals to shift out of the high wage into the low wage sector. This undoes some of the original increase in aggregate effort in the high wage sector and blunts the desirable effects of the original reduction in taxes.

Our paper also relates to earlier research on optimal income taxation in models with endogenous wages and occupational choice, such as Feldstein (1973), Zeckhauser (1977), Allen (1982), Boadway, Marceau, and Pestieau (1991), and Parker (1999). This literature has largely restricted attention to linear taxation. An exception is work by Moresi (1997), who considers non-linear taxation of profits in a model of occupational choice between workers and entrepreneurs. The occupational choice margin in his model is considerably simplified, however, and heterogeneity is confined to affect one occupation only, not the other.

Restricting heterogeneity to affect one occupation only, or tax schedules to be linear, sidesteps the complexities of multidimensional screening, which emerges naturally in the present model. In fact, few studies in the optimal taxation literature have attempted to deal with multidimensional screening problems until recently. In a recent study of the optimal income taxation of couples, Kleven, Kreiner and Saez (2009) have made progress along these lines, as have Choné and Laroque (2010). Both papers have significantly different information structures than ours, however. The second dimension of heterogeneity enters preferences additively in the former, and in the latter it is a taste for labor rather than a second standard skill type as we employ here.³

More generally, this paper builds on the large literature on optimal income taxation following the seminal contributions by Mirrlees (1971) and Diamond (1998). Until recently, the focus of the theoretical literature was on deriving results for a given assumed distribution of skills and social welfare function. Saez (2001) focused instead on inferring optimal taxes from observed income distributions. Moreover, Laroque (2005), Werning (2007) and Choné and Laroque (2010) study conditions under which an observer can test whether an existing set of taxes is or is not Pareto efficient. In the same spirit, we characterize the set of Pareto efficient tax policies rather than focusing on a particular social welfare function. With multiple complementary sectors, however, the wage distribution

³In Rothschild and Scheuer (2011), we use a similar model to the one considered here to characterize optimal corrective taxation in the presence of a rent-seeking sector.

is endogenous to the tax code, so earlier tests—e.g., Werning (2007), who infers wage-cum-skill distributions from income distributions as a test of optimality—are potentially misleading. One might conclude that the tax code is indeed Pareto efficient given the inferred skill distribution under the (implicit and incorrect) assumption that the skill distribution is independent of the tax code. Our concept of a self-confirming policy equilibrium, described above, is meant to capture this situation. It is closely related to the recent literature on self-confirming equilibria in learning models (e.g., Sargent, 2009, and Fudenberg and Levine, 2009).

This paper proceeds as follows. In section 2, we describe the basic model, and show that a single non-linear income tax implements the most general direct mechanism. In section 3, we characterize and compare the optimal and SCPE non-linear taxes. In Section 4, we point out the role of occupational choice and overlapping wage distributions by comparing our results to those in a discrete type model without occupational choice, as in Stiglitz (1982). Section 5 provides a stylized numerical example and Section 6 extends our results to an unbounded support of the skill distribution and additional heterogeneity in tastes for different occupations. Section 7 offers some brief conclusions. Proofs are in the Appendix.

2 The Model

2.1 Setup

We consider an economy where a unit mass of individuals can choose between working in either of two sectors. Accordingly, individuals have a two-dimensional skill type $(\theta, \varphi) \in \Theta \times \Phi$ with $\Theta = [\underline{\theta}, \bar{\theta}]$ and $\Phi = [\underline{\varphi}, \bar{\varphi}]$. θ captures an individual's productivity when working in the Θ -sector, whereas its skill in the Φ -sector is φ . These skills follow a distribution in the population described by the two-dimensional cdf $F(\theta, \varphi)$ and density $f(\theta, \varphi)$.

Individuals have preferences over consumption c and effort e captured by the utility function $U(c, e)$ with $U_c > 0$, $U_e < 0$. We denote the consumption, effort, utility, and sector assigned to an individual of type (θ, φ) by $c(\theta, \varphi)$, $e(\theta, \varphi)$, $V(\theta, \varphi) \equiv U(c(\theta, \varphi), e(\theta, \varphi))$, and $S(\theta, \varphi) \in \{\Theta, \Phi\}$, respectively.

The technology in the economy is described by a constant returns to scale (CRS) aggregate production function $Y(E_\theta, E_\varphi)$ that combines the skill-weighted aggregate effort in the two sectors to produce the consumption good. Formally, aggregate effort in the

two sectors is given by

$$E_\theta \equiv \int_{P \subset \Theta \times \Phi} \theta e(\theta, \varphi) dF(\theta, \varphi) \quad \text{and} \quad E_\varphi \equiv \int_{\Theta \times \Phi \setminus P} \varphi e(\theta, \varphi) dF(\theta, \varphi)$$

with

$$P \equiv \{(\theta, \varphi) | S(\theta, \varphi) = \Theta\}.$$

Since technology is linear homogeneous, the marginal products only depend on the ratio of aggregate effort in the two sectors $x \equiv E_\theta / E_\varphi$ and are therefore denoted by $Y_\theta(x)$ and $Y_\varphi(x)$. We define an individual's wage w as the marginal return to effort, so that

$$w = \begin{cases} Y_\theta(x)\theta & \text{if } S(\theta, \varphi) = \Theta, \\ Y_\varphi(x)\varphi & \text{if } S(\theta, \varphi) = \Phi. \end{cases} \quad (1)$$

An individual's income is then given by $y(\theta, \varphi) \equiv we(\theta, \varphi)$. As is standard, we assume that $U(c, e)$ satisfies the single-crossing property, i.e. an individual's marginal rate of substitution between income and consumption, $-U_e(c, y/w) / (wU_c(c, y/w))$ is decreasing.

2.2 Example: Entrepreneurs and Workers

Our model readily captures the situation where occupational choice is between becoming an entrepreneur and a worker, and entrepreneurs hire workers. To see this, suppose Θ is the entrepreneurial sector and Φ the workers' sector, and each individual has the according skills θ and φ for each activity. Moreover, suppose entrepreneurs combine their own effective entrepreneurial effort $\tilde{e}_\theta = \theta e$ and the effective labor supply of workers $\tilde{e}_\varphi = \varphi e$ in a constant returns to scale production function $\tilde{Y}$ to produce the consumption good. They hire labor in a competitive labor market, taking the wage $\tilde{w}_\varphi$ as given, so that their labor demand problem becomes, for any given $\tilde{e}_\theta$,

$$\pi(\tilde{e}_\theta) = \max_{\tilde{e}_\varphi} \tilde{Y}(\tilde{e}_\theta, \tilde{e}_\varphi) - \tilde{w}_\varphi \tilde{e}_\varphi.$$

Hence, they hire labor so as to equalize the marginal product of labor to the wage, i.e. $\tilde{Y}_\varphi = \tilde{w}_\varphi$ and, by constant returns to scale, profits are just

$$\pi(\tilde{e}_\theta) = \tilde{Y}_\theta \tilde{e}_\theta + \tilde{Y}_\varphi \tilde{e}_\varphi - \tilde{Y}_\varphi \tilde{e}_\varphi = \tilde{Y}_\theta \tilde{e}_\theta.$$

We can therefore view entrepreneurs as facing a wage $\tilde{w}_\theta = \tilde{Y}_\theta$ on their effective entrepreneurial effort $\tilde{e}_\theta$. Since all entrepreneurs have the same technology in terms of

effective effort, aggregating to total effective effort $E_\theta = \int \tilde{e}_\theta$ and $E_\varphi = \int \tilde{e}_\varphi$ implies $\tilde{w}_\theta = \tilde{Y}_\theta = Y_\theta(x)$ and $\tilde{w}_\varphi = \tilde{Y}_\varphi = Y_\varphi(x)$ as before.

2.3 A Mechanism Design Approach

We start with characterizing the implementation of a direct mechanism where individuals announce their privately known type (θ, φ) and then get assigned consumption $c(\theta, \varphi)$, income $y(\theta, \varphi)$ and a sector to work in $S(\theta, \varphi)$. Assuming that only income and consumption are observable, but an individual's sectoral choice is not, the resulting incentive constraints that guarantee truth-telling of the agents are as follows. First, suppose $S(\theta, \varphi) = \Theta$, i.e. we want to send type (θ, φ) to the Θ -sector. Then incentive compatibility requires that

$$U\left(c(\theta, \varphi), \frac{y(\theta, \varphi)}{Y_\theta(x)\theta}\right) \geq \max \left\{ U\left(c(\theta', \varphi'), \frac{y(\theta', \varphi')}{Y_\theta(x)\theta}\right), U\left(c(\theta', \varphi'), \frac{y(\theta', \varphi')}{Y_\varphi(x)\varphi}\right) \right\} \quad \forall \theta', \varphi'$$

since there are two ways for type (θ, φ) to imitate another type (θ', φ') , namely by earning (θ', φ') 's income either in the Θ - or the Φ -sector. Analogously, if $S(\theta, \varphi) = \Phi$, we need

$$U\left(c(\theta, \varphi), \frac{y(\theta, \varphi)}{Y_\varphi(x)\varphi}\right) \geq \max \left\{ U\left(c(\theta', \varphi'), \frac{y(\theta', \varphi')}{Y_\theta(x)\theta}\right), U\left(c(\theta', \varphi'), \frac{y(\theta', \varphi')}{Y_\varphi(x)\varphi}\right) \right\} \quad \forall \theta', \varphi'.$$

The following lemma shows that any incentive compatible direct mechanism in this framework can be implemented by offering a single (non-linear) income tax schedule.

Lemma 1. *Suppose that only income y and consumption c are observable, whereas an individual's skill type (θ, φ) , effort e and sectoral choice S are private information. Then any incentive compatible direct mechanism can be implemented by offering a schedule of $c(w), y(w)$ -bundles with $w \equiv \max\{Y_\theta(x)\theta, Y_\varphi(x)\varphi\}$ and*

$$S(\theta, \varphi) = \begin{cases} \Theta & \text{if } Y_\theta(x)\theta > Y_\varphi(x)\varphi, \\ \Phi & \text{if } Y_\theta(x)\theta < Y_\varphi(x)\varphi, \end{cases}$$

which is equivalent to offering a single (non-linear) income tax schedule.

Proof. In Appendix A.1. □

In other words, if an individual's sectoral choice is not observable, there is no loss in considering the set of allocations implemented by offering a single non-linear income tax schedule and letting individuals choose their preferred income and sector given this.

In particular, any such allocation will treat any two individuals who earn the same wage (albeit in different sectors) the same, and individuals will work in the sector in which they can achieve a higher wage. We study these allocations below.

Even though we do not consider it in (this version of) this paper, it is worth observing that a similar analysis applies if sectoral choice is observable, so that the social planner can condition taxes on sectoral choice. In this case, any incentive compatible allocation can be implemented by offering two non-linear income tax schedules—one for each sector.

3 Characterizing Optimal Income Taxes

By Lemma 1, when the income tax does not condition on sectoral choice, individuals will choose sectors based only on their wage $w = \max\{Y_\theta(x)\theta, Y_\varphi(x)\varphi\}$ and all individuals with the same wage get the same allocation $c(w), y(w)$ or, equivalently, $V(w), e(w)$. (With mild notational abuse, we will use $c(V, e)$ to denote the inverse of $U(c, e)$ with respect to the first argument.)

3.1 Pareto Optima and Self-Confirming Equilibria

We start with observing that, for any given x , the marginal productivities in the two sectors $Y_\theta(x)$ and $Y_\varphi(x)$, the two-dimensional skill distribution $F(\theta, \varphi)$ and the implied sectoral choice induce a one-dimensional wage distribution characterized by the cdf

$$F_x(w) \equiv F\left(\frac{w}{Y_\theta(x)}, \frac{w}{Y_\varphi(x)}\right)$$

and the sectoral densities

$$f_x^\theta(w) = \frac{1}{Y_\theta(x)} \int_{\underline{\varphi}}^{w/Y_\varphi(x)} f\left(\frac{w}{Y_\theta(x)}, \varphi\right) d\varphi, \quad f_x^\varphi(w) = \frac{1}{Y_\varphi(x)} \int_{\underline{\theta}}^{w/Y_\theta(x)} f\left(\theta, \frac{w}{Y_\varphi(x)}\right) d\theta$$

with associated cdfs $F_x^\theta(w)$ and $F_x^\varphi(w)$ and with $f_x(w) = f_x^\theta(w) + f_x^\varphi(w)$. We also define the bottom and top wages given x as

$$\underline{w}_x = \max\{Y_\theta(x)\underline{\theta}, Y_\varphi(x)\underline{\varphi}\}, \quad \bar{w}_x = \max\{Y_\theta(x)\bar{\theta}, Y_\varphi(x)\bar{\varphi}\}.$$

If we assign Pareto weights $G(\theta, \varphi)$ in the two-dimensional skill space, we analogously obtain a distribution of Pareto weights over wages given x , denoted by $G_x(w)$, with the corresponding densities $g_x(w) = g_x^\theta(w) + g_x^\varphi(w)$ (and cdfs $G_x^\theta(w)$ and $G_x^\varphi(w)$). Noting

that

$$x = \frac{\int_{\underline{w}_x}^{\bar{w}_x} \frac{w}{Y_\theta(x)} e(w) f_x^\theta(w) dw}{\int_{\underline{w}_x}^{\bar{w}_x} \frac{w}{Y_\varphi(x)} e(w) f_x^\varphi(w) dw} \equiv \tilde{x}(x), \quad (2)$$

we can therefore write the Pareto problem for income taxation in the Roy model as

$$\max_{x, e(w), V(w)} \int_{\underline{w}_x}^{\bar{w}_x} V(w) dG_x(w) \quad (3)$$

s.t.

$$V'(w) + U_e(c(V(w), e(w)), e(w)) \frac{e(w)}{w} = 0 \quad \forall w \in [\underline{w}_x, \bar{w}_x] \quad (4)$$

$$\Gamma(x) \int_{\underline{w}_x}^{\bar{w}_x} w e(w) f_x^\theta(w) dw - \int_{\underline{w}_x}^{\bar{w}_x} w e(w) f_x^\varphi(w) dw = 0 \quad (5)$$

with $\Gamma(x) \equiv Y_\varphi(x) / x Y_\theta(x)$ and

$$\int_{\underline{w}_x}^{\bar{w}_x} (w e(w) - c(V(w), e(w))) f_x(w) dw \geq 0. \quad (6)$$

We refer to the three constraints (4), (5) and (6) as the incentive constraints, the consistency condition and the resource constraint, respectively.⁴

We can decompose the Pareto problem (3) to (6) into an inner problem, given x , and an outer problem, which maximizes over x . Formally, fix some x and let $W(x)$ denote the value of the objective (3) when maximizing it over $V(w), e(w)$ subject to (4), (5) and (6) (the inner problem). Then the outer problem is simply $\max_x W(x)$.⁵

For some of the subsequent analysis, it will be useful to restrict attention to the subset of allocations on the Pareto frontier that result from attaching the same Pareto weight to any two individuals who end up earning the same wage (and thus income), regardless of their sectoral choice. These allocations are obtained by solving the Pareto problem using relative welfare weights $\Psi(F_x(w))$ rather than the more general Pareto weights $G_x(w)$. Intuitively, $\Psi(\cdot)$ determines how much social welfare weight is attached to different quan-

⁴Note that we have made use of the local version of the incentive constraints in (4) and dropped the additional monotonicity constraint that $y(w) = w e(w)$ must be increasing in w (which would guarantee global incentive compatibility together with (4) and the single-crossing condition). We thus abstract from issues of bunching.

⁵Because $W(x)$ may not be concave, there may be Pareto optimal allocations that do not solve a Pareto problem for any Pareto weights. The first order conditions for some Pareto problem *will* be satisfied at any Pareto optimum, however.

tiles of the wage (and thus income) distribution.⁶ Lemma 2 shows that using the relative Pareto weights $\Psi(\cdot)$ indeed produces Pareto optimal allocations.

Lemma 2. *Any solution $\{x, e(\theta, \varphi), V(\theta, \varphi)\}$ to the planning problem (3) to (6) with relative welfare weights $\Psi(F_x(w))$ in (3) is Pareto optimal.*

Proof. See Appendix B.1 □

To compare Pareto optimal tax schedules to those that would be optimal in a standard Mirrlees model with an exogenous skill and thus wage distribution, we consider a naive social planner who incorrectly presumes that the wage distribution is exogenous. We define a self-confirming equilibrium (SCPE) as a mutually consistent tax schedule and wage distribution pair where, when the planner designs an optimal tax schedule taking the wage distribution as given, this tax policy induces the wage distribution that was taken as given.⁷ In particular, if the government identifies the wage distribution from the observed income distribution and for a given tax schedule, as suggested by Saez (2001), and designs an optimal tax schedule taking this wage distribution as given, then the optimality of the tax schedule is confirmed in a SCPE.

Formally, and using relative welfare weights Ψ , the inner SCPE problem for a given x is just

$$\max_{x, e(w), V(w)} \int_{\underline{w}_x}^{\bar{w}_x} V(w) d\Psi(F_x(w)) \quad (7)$$

subject to (4) and (6). The consistency constraint (5) disappears because such a planner is not aware of the fact that the wage distribution is endogenous and the allocation has to be consistent with x . But to ensure that the allocation in aggregate is consistent, we have to require that the value of x implied by the solution to the inner problem equals x . Formally, the outer problem requires finding a fixed point of the mapping $x \rightarrow \tilde{x}^*(x)$, where $\tilde{x}^*(x)$ is defined by (2), evaluated at the $e(w), V(w)$ that solves the inner SCPE problem. To summarize, we define a SCPE as follows:

Definition 1. *A self-confirming equilibrium (SCPE) is a value of x and an allocation $V(w), e(w)$ such that (i) x is a fixed point of $\tilde{x}^*(x)$ and (ii) $V(w)$ and $e(w)$ solve the inner SCPE problem given x .*

⁶These allocations would be obtained by considering Pareto weights in the (θ, φ) -space of the form $G(\theta, \varphi) = \Psi(F(\theta, \varphi))$, so that $G_x(w) = \Psi(F_x(w))$.

⁷This definition is analogous to the notion of SCPE introduced in Rothschild and Scheuer (2011) in the context of rent-seeking.

3.2 Inner Problem

We can use the inner problem, for fixed x , to derive formulas for marginal tax rates. We start with Pareto optimal tax schedules using general Pareto weights $G(\theta, \varphi)$.

Proposition 1. *Let μ denote the multiplier on the resource constraint (6), let $\mu\zeta$ denote the multiplier on the consistency condition (5), and use $\varepsilon^u(w)$ and $\varepsilon^c(w)$ to denote the uncompensated and compensated labor supply elasticities, respectively. For given x and Pareto weights G , marginal tax rates satisfy*

$$1 - T'(y(w)) = \frac{1 + \zeta(1 + \Gamma(x)) \left(\frac{Y_\varphi(x)}{Y_\varphi(x) + xY_\theta(x)} - \frac{f_x^\varphi(w)}{f_x(w)} \right)}{1 + \frac{\eta(w)}{wf_x(w)} \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)}} \quad (8)$$

where

$$\eta(w) = \int_w^{\bar{w}_x} \left(1 - \frac{g_x(z)}{f_x(z)} \frac{U_c(z)}{\mu} \right) \exp \left(\int_w^z \left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)} \right) \frac{dy(s)}{y(s)} \right) f_x(z) dz. \quad (9)$$

Proof. See Appendix B.2. □

Since the inner SCPE problem differs from the inner Pareto problem because of the absence of the consistency condition (5), the marginal tax rates for the SCPE can be found by using relative Pareto weights Ψ and setting $\zeta = 0$ in Proposition 1, as in the following Corollary.

Corollary 1. *For given x and relative Pareto weights Ψ , marginal tax rates satisfy*

$$1 - T'(y(w)) = \frac{1}{1 + \frac{\eta(w)}{wf_x(w)} \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)}} \quad (10)$$

in an SCPE and

$$1 - T'(y(w)) = \frac{1 + \zeta(1 + \Gamma(x)) \left(\frac{Y_\varphi(x)}{Y_\varphi(x) + xY_\theta(x)} - \frac{f_x^\varphi(w)}{f_x(w)} \right)}{1 + \frac{\eta(w)}{wf_x(w)} \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)}} \quad (11)$$

in a Pareto optimum, where

$$\eta(w) = \int_w^{\bar{w}_x} \left(1 - \psi(F_x(z)) \frac{U_c(z)}{\mu}\right) \exp\left(\int_w^z \left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)}\right) \frac{dy(s)}{y(s)}\right) f_x(z) dz. \quad (12)$$

These formulas show that the formula for marginal keep shares $1 - T'$ is adjusted in the Pareto problem compared to the SCPE problem by a correction factor that depends on ξ and a comparison between the aggregate income share of the Φ -sector, given by

$$\frac{Y_\varphi(x)}{Y_\varphi(x) + xY_\theta(x)} = \frac{Y_\varphi(x)E_\varphi}{Y_\varphi(x)E_\varphi + Y_\theta(x)E_\theta} = \frac{Y_\varphi(x)E_\varphi}{Y(E_\theta, E_\varphi)},$$

with its local income share $y(w)f_x^\varphi(w)/(y(w)f_x(w)) = f_x^\varphi(w)/f_x(w)$.⁸

This is intuitive. For instance, suppose $\xi > 0$ (we will show below that this corresponds to the case where Φ is the low-income sector). Then the marginal keep share is scaled down in the Pareto problem relative to the SCPE whenever, at the given wage (or equivalently income) level, the local income share of the Φ -sector exceeds its aggregate income share. This disproportionately discourages Φ -sector effort and therefore raises wages in the Φ -sector relative to the Θ -sector. Hence, the solution to the Pareto problem uses this “indirect redistribution” channel through wages in order to redistribute to the low-income sector, which is desirable for relative Pareto weights with $\Psi(F) \geq F$ for all $F \in [0, 1]$. Note that this implies a force towards less progressivity in the Pareto problem relative to the SCPE: If Φ is the low-income sector, marginal tax rates will be scaled up in the Pareto problem compared to the SCPE for low income levels, and scaled down for high income levels.⁹

Note that, in particular, the top marginal tax rate is not generally zero in a Pareto optimum. It is given by

⁸It is important, of course, to keep in mind that the formulas in the preceding corollary are evaluated at endogenously determined values of x . Since, in general, the level of x in the SCPE and the solution to the Pareto problem will differ for a given economy, the formulas do not permit a direct comparison of tax rates at the two solutions. One interpretation of such a comparison is as a comparison of the tax rates in two *different* economies that endogenously happen to have the same skill distribution. By comparing the tax rates with the two formulas, one can infer whether the social planner is solving for an SCPE or a Pareto optimum.

⁹The same would be true if Φ were the high-income sector. In this case, we will show below that $\xi < 0$ and hence marginal tax rates are again higher in the Pareto optimum compared to the SCPE for low income levels, and lower for high income levels, according to (11).

Corollary 2. *The top marginal tax rate is zero in any SCPE and given by*

$$T'(y(\bar{w}_x)) = \xi(1 + \Gamma(x)) \left(\frac{f_x^\varphi(\bar{w}_x)}{f_x(\bar{w}_x)} - \frac{Y_\varphi(x)}{Y_\varphi(x) + xY_\theta(x)} \right)$$

in any Pareto optimum. In particular, if $f_x^\varphi(\bar{w}_x)/f_x(\bar{w}_x) = 1$ (respectively $f_x^\varphi(\bar{w}_x)/f_x(\bar{w}_x) = 0$), then $T'(y(\bar{w}_x)) = \xi$ (respectively $T'(y(\bar{w}_x)) = -\xi\Gamma(x)$).

It will turn out below that $\xi > 0$ if Φ is the low-income sector, and vice versa. But if Φ is the low-income sector, its local income share at the top of the income distribution is smaller than its aggregate income share, and vice versa. Hence, by Corollary 2, the top marginal tax rate will generally be negative in a Roy model with regular Pareto weights. For instance, consider the special case of a constant elasticity of substitution (CES) production function

$$Y(E_\theta, E_\varphi) = [\alpha E_\theta^\rho + (1 - \alpha) E_\varphi^\rho]^{1/\rho}, \quad (13)$$

where $1/(1 - \rho)$ is the constant elasticity of substitution. Then the top marginal tax rate in the solution to the Pareto problem simplifies as follows:

Corollary 3. *With CES technology, the Pareto optimal top marginal tax rate is*

$$T'(y(\bar{w}_x)) = \xi \left(1 + \frac{1 - \alpha}{\alpha} x^{-\rho} \right) \left(\frac{f_x^\varphi(\bar{w}_x)}{f_x(\bar{w}_x)} - \frac{1 - \alpha}{\alpha x^\rho + 1 - \alpha} \right)$$

and with Cobb-Douglas technology ($\rho = 0$),

$$T'(y(\bar{w}_x)) = \frac{\xi}{\alpha} \left(\frac{f_x^\varphi(\bar{w}_x)}{f_x(\bar{w}_x)} - (1 - \alpha) \right).$$

With Cobb-Douglas technology, the aggregate income share of the Φ -sector is independent of x and given by the constant $1 - \alpha$, so that we only need to (i) compare the local income (or population) share of the Φ -sector workers among the top wage earners to $1 - \alpha$ and (ii) determine the sign of the multiplier ξ in order to sign the top marginal tax rate. In order to achieve the second requirement, we consider the outer problem in the following section.

3.3 Outer Problem

For the following, the substitution elasticity of the production function $Y(E_\theta, E_\varphi)$, denoted by $\sigma(x)$, will be useful. The following Lemma provides a simple expression for it.

Lemma 3. *The substitution elasticity of $Y(E_\theta, E_\varphi)$ is given by*

$$\sigma(x) \equiv -\frac{1}{x\lambda(x)} \quad \text{with} \quad \lambda(x) \equiv \frac{Y'_\theta(x)}{Y_\theta(x)} - \frac{Y'_\varphi(x)}{Y_\varphi(x)}.$$

Proof. See Appendix B.3 □

Using this, we can derive the following decomposition of the welfare effect of a marginal change in x .

Lemma 4. *For any Pareto weights G , the welfare effect of a marginal change in x can be decomposed as follows:*

$$W'(x) = -\frac{1}{x\sigma(x)} (I + R + \xi\mu [Y_\varphi(x)E_\varphi\sigma(x) + (1 + \Gamma(x))(S + C)]), \quad (14)$$

where

$$S \equiv \frac{1}{Y_\theta(x)Y_\varphi(x)} \int_{\underline{w}_x}^{\bar{w}_x} w^2 e(w) f\left(\frac{w}{Y_\theta(x)}, \frac{w}{Y_\varphi(x)}\right) dw > 0, \quad (15)$$

$$I \equiv \mu \int_{\underline{w}_x}^{\bar{w}_x} \eta(w) w \frac{V'(w)}{U_c(w)} \frac{d}{dw} \left(\frac{f_x^\varphi(w)}{f_x(w)} \right) dw \quad (16)$$

$$R \equiv \int_{\underline{w}_x}^{\bar{w}_x} w V'(w) \frac{f_x^\theta(w) f_x^\varphi(w)}{f_x(w)} \left(\frac{g_x^\theta(w)}{f_x^\theta(w)} - \frac{g_x^\varphi(w)}{f_x^\varphi(w)} \right) dw, \quad (17)$$

and

$$C \equiv \int_{\underline{w}_x}^{\bar{w}_x} w^2 e'(w) \frac{f_x^\varphi(w) f_x^\theta(w)}{f_x(w)} dw. \quad (18)$$

Proof. In Appendix B.4. □

We provide a heuristic derivation that reveals the intuition behind this result. To that end, we break the welfare effects of a small change in x into four effects:

1. A *direct* effect: the effect of changing x on $\Gamma(x)$ in (5), holding wages and sectors of each individual constant.
2. A *direct wage shift* effect: the effect arising from the change in wages induced by the change in x , holding each individual's allocation and sector constant.
3. An *indirect wage shift* effect: the effect arising from the change in allocations caused by the change in wages induced by the change in x , holding the individual's sector constant.

4. A *sectoral shift* effect: the effect that arises from individuals changing sectors, holding their wages constant.

Notice first that if technology is linear, so that $\sigma(x) = \infty$, then setting $W'(x) = 0$ immediately implies $\xi = 0$. By Corollary 1, the marginal tax formulas for the SCPE and Pareto problems coincide. This is intuitive: in this case, wages are exogenous to the tax code, so the fact that there are two sectors is irrelevant. It is only the wage shift and sectoral shift effects driven by the endogeneity of wages in the finite $\sigma(x)$ case that provide scope for using additional tools for accomplishing redistributive objectives.

Because individual allocations are held constant, the direct wage shift has no effect on the objective. Because of constant returns to scale, it also has no effect on the resource constraint. However, it affects both the incentive and consistency constraints. The effect on the latter can be combined with the direct effect (of changing x on $\Gamma(x)$) to yield $-Y_\varphi(x)E_\varphi/x$. To wit: re-write the consistency condition (5) as $Y_\varphi(x)(E_\theta/x - E_\varphi) = 0$, and consider the effect of a change in x , holding sectors and efforts constant. The terms E_θ and E_φ are unchanged, so the effect is just $-Y_\varphi(x)E_\theta/x^2 = -Y_\varphi(x)E_\varphi/x$.

The term in expression (14) containing I arises from the direct effect of the wage shift on the incentive constraints. To understand it, consider the effects of a small decrease in x for a portion of the wage distribution centered at wage w . Such a decrease will raise Θ -sector wages and lower Φ -sector wages. If the share of Φ -sector workers is *increasing* locally, this leads to a local compression of wage distribution. Such a compression eases the incentive compatibility constraints if they are binding in the downward direction (i.e., higher wage individuals need to be prevented from imitating lower wage individuals)—i.e., if $\eta(w) > 0$. A decrease in x therefore leads to a welfare improvement insofar as $\eta(w)d(f_x^\varphi(w)/f_x(w))/dw > 0$ (and the magnitude of this improvement will be related to how steeply increasing the utility distribution is). As we will formalize in the subsequent section, I can be thought of as a (generalized) Stiglitz (1982) effect: with endogenous wages, increasing (decreasing) effort at high (low) wages will raise (lower) wages at low (high) wages.

The sectoral shift and indirect wage-shift effects are effects that are *not* present in Stiglitz's (1982) framework. It is therefore worth elaborating on why they take the form they do, and in particular, why they reinforce each other whenever $e'(w) > 0$. We consider the sectoral shift effect first.

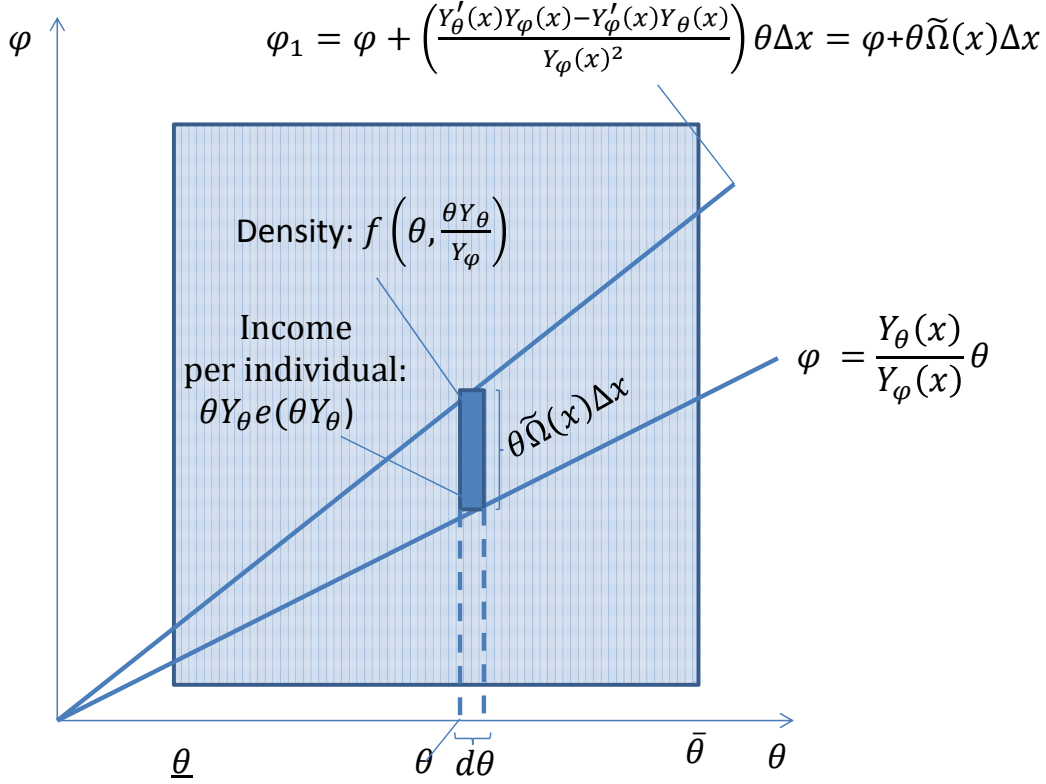


Figure 1: Illustrating the computation of the sectoral shift effect

3.3.1 The sectoral shift

To compute the sectoral shift effect, it is useful to write the consistency condition as $\Gamma(x)I_\theta(x) - I_\varphi(x) = 0$, where $I_\theta(x) \equiv \int_{\underline{w}}^{\bar{w}} we(w)f_x^\theta(w)dw$ is the income earned in the Θ -sector, and similarly for $I_\varphi(x)$. Consider a small decrease Δx in x , holding efforts and wages constant. This will lead some individuals to shift from the Φ - to the Θ -sector, as illustrated in figure 1. Let $\Delta I_\theta(x)$ denote the resulting change in Θ -sector income. Since there is an equal and opposite change in Φ -sector income, the sectoral shift effect can be written as

$$S = (\Gamma(x) + 1)\Delta I_\theta(x).$$

Figure 1 illustrates the computation of $\Delta I_\theta(x)$. It considers the mass element of individuals with Θ -sector skills between θ and $\theta + d\theta$ who are in the Φ -sector at x but the Θ -sector at $x + \Delta x$. The height of this element is

$$\frac{d(Y_\theta(x)/Y_\varphi(x))}{dx} \theta \Delta x = \left(\frac{Y'_\theta(x) Y_\varphi(x) - Y_\theta(x) Y'_\varphi(x)}{Y_\varphi(x)^2} \right) \theta \Delta x.$$

The income earned by each individual in that element is $\theta Y_\theta(x) e(\theta Y_\theta(x))$, and the density of individuals is $f\left(\theta, \theta \frac{Y_\theta(x)}{Y_\varphi(x)}\right)$. Multiplying the width ($d\theta$) by the height, the density, and the per individual income, and then integrating over θ gives:

$$\Delta I_\theta(x) = \Delta x \int_{\underline{\theta}}^{\bar{\theta}} \left(\frac{Y'_\theta(x) Y_\varphi(x) - Y_\theta(x) Y'_\varphi(x)}{Y_\varphi(x)^2} \right) Y_\theta(x) \theta^2 e(\theta Y_\theta(x)) f\left(\theta, \frac{\theta Y_\theta(x)}{Y_\varphi(x)}\right) d\theta. \quad (19)$$

Changing variables to $w = \theta Y_\theta(x)$ yields,

$$\Delta I_\theta(x) = \Delta x \frac{\lambda(x)}{Y_\theta(x) Y_\varphi(x)} \int_{\underline{w}}^{\bar{w}} w^2 e(w) f\left(\frac{w}{Y_\theta(x)}, \frac{w}{Y_\varphi(x)}\right) dw. \quad (20)$$

and the sectoral shift term S defined in expression (15) follows directly.

3.3.2 The indirect wage shift effect

There is no simple graphical representation of the indirect wage shift effect, but a similar heuristic could be used to derive the terms C and R . We omit the algebraic details here, and instead providing the basic intuition behind those effects. (See Appendix B.4 for a formal treatment.)

Imagine increasing x by a small amount while holding the tax code constant so that allocations $e(w)$ and $V(w)$ are an unchanged function of wages. The two types of individuals at original wage w^* are affected differently by the change in x : individuals in the Φ -sector find their (Φ -sector) wage increases; Θ -sector individuals see their wage decrease. The Θ -sector individuals move down along the (fixed) schedules $e(w)$ and $V(w)$, and Φ -sector individuals move up. Depending on the schedules and the proportions of the two types at w^* , the net effort and utility effect on wage w^* -individuals may be positive or negative. The algebraic manipulations in Appendix B.4 are motivated by thinking of these shifts in two steps: a level shift of all wage w^* -individuals that absorbs the *net* shift of effort and utility, and a *re-allocation* of effort and utility across the two types. The former involves a particular shift in the $e(w)$ and $V(w)$ schedules, and has zero welfare effects by an envelope argument (the original schedule was optimal). The re-allocation of effort and utility in the latter respectively give rise to the terms C and R in Lemma 4.

Because it arises from a re-allocation of utility across individuals at the same w^* induced by the change in x , the term R disappears when there is no intrinsic sectoral preference—i.e., when $g_x^\theta(w)/g_x^\varphi(w) = f_x^\theta(w)/f_x^\varphi(w)$ for all w . It is straightforward to show that this will be the case whenever $G(\theta, \varphi)$ takes the form $G(\theta, \varphi) = \Psi(F(\theta, \varphi))$ or with relative welfare weights $\Psi(F_x(w))$. In contrast, when the social planner has an in-

trinsic preference for the Θ -sector individuals at wage w^* , the re-allocation of utility from the Θ - to the Φ -sector is welfare reducing.

The term C arises from an analogous re-allocation of effort. If the effort schedule is increasing (i.e. $e'(w) \geq 0$), then an increase in x effectively re-allocates effort from the Θ - to the Φ -sector, because Θ -workers move down and Φ -workers up along the $e(w)$ schedule. This effect therefore reinforces the sectoral shift effect.

3.4 Marginal Tax Rate Results

We can use the decomposition in Lemma 4 to sign the multiplier on the consistency condition ξ at an optimal x by setting $W'(x) = 0$:

$$\xi = \frac{-(I + R)/\sigma(x)}{Y_\varphi(x)E_\varphi + \frac{1 + \Gamma(x)}{\sigma(x)}(C + S)}. \quad (21)$$

We summarize the resulting conditions for the sign of ξ in the following corollary:

Corollary 4. *With linear technology ($\sigma(x) = \infty$), $\xi = 0$.*

For $\sigma(x) \in (0, \infty)$, the following holds for any Pareto optimum with (i) increasing effort ($e'(w) \geq 0$) and (ii) downwards-binding incentive constraints ($\eta(w) \geq 0$ for all w):

1. $\xi \geq 0$ if $f_x^\theta(w)/f_x(w)$ is increasing in w and $g_x^\theta(w)/f_x^\theta(w) \leq g_x^\varphi(w)/f_x^\varphi(w) \forall w$,
2. $\xi \leq 0$ if $f_x^\varphi(w)/f_x(w)$ is increasing in w and $g_x^\theta(w)/f_x^\theta(w) \geq g_x^\varphi(w)/f_x^\varphi(w) \forall w$.

The inequalities in (1) and (2) are strict if $\eta(w)$ is not identically zero.

Conditions (i) and (ii) are sufficient, but not necessary conditions. The former ensures that the indirect wage shift term C reinforces the sectoral shift effect. The latter holds whenever the (average) marginal social value of consumption, given by $U_c(w)g_x(w)/f_x(w)$, is decreasing. This is guaranteed with quasilinear-in-consumption preferences and weakly progressive welfare weights (such that $g_x(w)/f_x(w)$ is increasing), for example. It ensures that a compression of the wage distribution eases the incentive compatibility constraints. If $f_x^\theta(w)/f_x(w)$ is increasing in w , then Θ is the high-skill sector, and an increase in x , by raising wages in the Φ -sector and lowering them in the Θ -sector, has desirable wage compression effects, as in Stiglitz (1982). This desirable effect is reinforced by R whenever $g_x^\varphi(w)/f_x^\varphi(w) \geq g_x^\theta(w)/f_x^\theta(w) \forall w$. In this case, the social planner puts higher social welfare weight on Φ -sector workers than on Θ -sector workers at any given wage, and the wage changes induced by an increase in x also have direct benefits.

Combining these results from the outer problem with the marginal tax rate results from the inner problem has crisp implications for the comparison between Pareto optimal and SCPE tax schedules. For instance, suppose Θ is the high-skilled sector, i.e. $f_x^\varphi(w)/f_x(w)$ is decreasing so that $\xi > 0$ by Corollary 4. Then the marginal keep share in the Pareto optimum is scaled down relative to the SCPE if the local income share in the Φ -sector is higher than in aggregate. This disproportionately reduces Φ -sector effort and therefore indirectly increases wages in the Φ -sector, achieving redistribution to the low-skilled sector. In particular, since $f_x^\varphi(w)/f_x(w)$ is decreasing, this means that marginal keep shares are scaled down for low wages and scaled up for high wages and the top marginal tax rate is negative.

On the other hand, suppose $f_x^\varphi(w)/f_x(w)$ is increasing, so that $\xi < 0$. Marginal keep shares will be scaled down whenever $f_x^\varphi(w)/f_x(w)$ is low, i.e. again for high wages. The top marginal tax rate is also again negative. I.e. in *both* of the two cases the general equilibrium effects in the Roy model work in favor of *less* progressive taxation. We summarize these insights in the following Proposition:

Proposition 2. *If $\sigma(x) \in (0, \infty)$, then the top marginal tax rate is negative in any Pareto optimum with*

- (a) *a decreasing i -sector share of workers $f_x^i(w)/f_x(w)$, $i \in \{\theta, \varphi\}$,*
- (b) *an increasing effort schedule $e(w)$,*
- (c) *a decreasing social marginal utility of consumption schedule $U_c(w)g_x(w)/f_x(w)$ and*
- (d) *a weak intrinsic social preference for the i -sector, i.e. $g_x^i(w)/f_x^i(w) \geq g_x^j(w)/f_x^j(w)$ for all $w, j \neq i \in \{\theta, \varphi\}$.*

Notably, consider the special case with relative welfare weights $G_x(w) = \Psi(F_x(w))$. Then $g_x(w) = \psi(F_x(w))f_x(w)$ and hence

$$g_x^\theta(w) = \psi(F_x(w))f_x^\theta(w) \text{ and } g_x^\varphi(w) = \psi(F_x(w))f_x^\varphi(w).$$

This immediately implies $g_x^\theta(w)/f_x^\theta(w) = g_x^\varphi(w)/f_x^\varphi(w) \forall w$ and thus $R = 0$. With relative welfare weights, condition (d) can therefore be dropped.

Hence, these results reveal the following intuitive separation: Per Corollary 4, the sign of the multiplier ξ on the consistency constraint accounts for the overall redistributive motive across sectors, i.e. whether we want to redistribute from Θ to Φ or vice versa. Then conditional on this direction, the nonlinear marginal tax rate correction in the Pareto

optimum relative to the SCPE is determined by comparing local and aggregate income shares between sectors, per Proposition 1.

4 The Role of Occupational Choice and Continuous Types

In this section, we relate our results to those in Stiglitz (1982), who considers optimal nonlinear taxation in a two-type model with endogenous wages, but without occupational choice. We demonstrate that the general Roy model, with continuous types and occupational choice, features three extra effects, as captured by S , C and R in the previous section, that do not appear in any generic discrete type model. The disappearance of the sectoral shift effect S in a model without occupational choice is obvious. In addition, the Roy model with continuous types generates overlapping wage distributions in the two sectors, which gives rise to the effects C and R . In contrast, in a discrete type model, generically – and somewhat unrealistically – there are no workers in different sectors earning the same wage.

Moreover, we show that the extra “Roy” effects that emerge in our model do not change the sign of the general equilibrium effects found in Stiglitz (1982), but they mitigate them. In this sense, redistribution in the Roy model, while implying a less progressive optimal tax schedule than a standard Mirrlees model (as captured by a SCPE), leads to *more* progressive taxes than a discrete type model without occupational choice.

We start by formulating Stiglitz’s (1982) model in terms of the decomposition into an inner problem (for fixed x) and outer problem (optimizing over x) as so far. Let there be two types with skills θ and φ and with fractions f_θ and $f_\varphi = 1 - f_\theta$ in the population. We put (relative) Pareto weights ψ_θ and ψ_φ on them such that $f_\theta\psi_\theta + f_\varphi\psi_\varphi = 1$. Without loss of generality, we will think of θ as the high wage sector and φ as the low wage sector, so that regular welfare weights satisfy $\psi_\theta \leq 1$ and $\psi_\varphi \geq 1$. As in Stiglitz (1982), we therefore focus on the case where only the θ -type’s incentive constraint binds.

4.1 Inner Problem

Individuals are paid their marginal products, $w_\theta = \theta Y_\theta(x)$, and $w_\varphi = \varphi Y_\varphi(x)$. Hence, we can write the inner problem for fixed x as

$$W(x) = \max_{e_\theta, e_\varphi, V_\theta, V_\varphi} f_\theta\psi_\theta V_\theta + f_\varphi\psi_\varphi V_\varphi \quad (22)$$

subject to

$$V_\theta \geq U \left(c(c_\varphi, V_\varphi), e_\varphi \frac{w_\varphi}{w_\theta} \right), \quad (23)$$

$$\Gamma(x) f_\theta w_\theta e_\theta = f_\varphi w_\varphi e_\varphi, \quad (24)$$

$$f_\theta w_\theta e_\theta + f_\varphi w_\varphi e_\varphi \geq f_\theta c(V_\theta, e_\theta) + f_\varphi c(V_\varphi, e_\varphi). \quad (25)$$

As before, the outer problem is just $\max_x W(x)$.

We focus on the top marginal tax rate and consider the optimal allocation for the θ -type. Denoting by μ the multiplier on (25) and $\xi\mu$ on (24), the first order condition w.r.t. e_θ is

$$-f_\theta \mu (\partial c(V_\theta, e_\theta) / \partial e_\theta - w_\theta) + f_\theta \Gamma(x) \xi \mu w_\theta = 0. \quad (26)$$

Using $\partial c / \partial e = -U_e / U_c = MRS$, this simplifies to $MRS_\theta = w_\theta(1 + \Gamma(x)\xi)$. By the first order condition for the worker's utility maximization problem, i.e., $MRS/w = 1 - T'(y)$, this implies that the marginal tax rate for the high-wage, Θ -sector individual, is $-\Gamma(x)\xi$ as in Corollary 2.

4.2 Outer Problem

We next turn to the outer problem to determine ξ . As before, we can decompose the welfare effect of a marginal change in x into a direct effect (the derivative with respect to x , holding wages constant), and wage shift effects (from the direct and indirect effects of changes in w_θ and w_φ). The direct effect is simply

$$\xi \mu \Gamma'(x) f_\theta w_\theta e_\theta = \xi \mu \Gamma'(x) Y_\theta(x) E_\theta. \quad (27)$$

We can break the wage shift effects into the effect on the objective and the effects on each of the 3 constraints. By the envelope theorem, we can hold V_i and e_i , $i = \theta, \varphi$, constant. This makes the effect on the objective identically zero. The effect on the resource constraint (25) is also zero, since

$$f_\theta e_\theta \frac{dw_\theta}{dx} + f_\varphi e_\varphi \frac{dw_\varphi}{dx} = E_\theta Y'_\theta(x) + E_\varphi Y'_\varphi(x) = 0$$

by constant returns to scale. To compute the effect on (24), re-write the constraint as

$$\Gamma(x) (f_\theta w_\theta e_\theta + f_\varphi w_\varphi e_\varphi) - (1 + \Gamma(x)) f_\varphi w_\varphi e_\varphi,$$

and observe that the derivative of the first term is zero, by the preceding argument. Hence, the effect here is

$$-\xi\mu(1+\Gamma(x))E_\varphi Y'_\varphi(x). \quad (28)$$

Combining (27) and (28) yields $-\xi\mu Y_\varphi(x)E_\varphi/x$ just as in section 3.3.

Finally, putting a multiplier $\hat{\eta}\mu$ on (23), the effect of the wage shift on the incentive constraint is

$$\begin{aligned} & -\mu\hat{\eta}U_e\left(c_\varphi, e_\varphi \frac{w_\varphi}{w_\theta}\right) e_\varphi \frac{\varphi}{\theta} \left[\frac{Y'_\varphi(x)}{Y_\theta(x)} - \frac{Y_\varphi(x)Y'_\theta(x)}{Y_\theta(x)Y_\theta(x)} \right] \\ &= \lambda(x)\mu\hat{\eta}U_e\left(c_\varphi, e_\varphi \frac{w_\varphi}{w_\theta}\right) e_\varphi \frac{w_\varphi}{w_\theta} \equiv \lambda(x)\hat{I}. \end{aligned} \quad (29)$$

To see why calling this effect $\hat{I}$ by analogy to I from the general Roy model is appropriate, consider the definition of I in equation (16) and take the limit where $f_x^\theta(w)/f_x(w)$ is 0 up until w_θ , and 1 thereafter. Then the ratio $f_x^\theta(w)/f_x(w)$ is a step function, the derivative of which is the Dirac δ -function. Hence, the integral in (16) evaluates to $-V'(w_\theta)w_\theta\eta(w_\theta) = U_e(c_\theta, e_\theta)e_\theta\eta(w_\theta)$ by the incentive constraint (4). The only difference from $\hat{I}$ is that it has $e_\varphi w_\varphi/w_\theta$ instead of e_θ and c_φ rather than c_θ (and $\hat{\eta} = \eta/U_c$ is discrete rather than continuous). This difference is a result of the fact that the incentive constraint is discrete rather than local: in the density limit, the θ -type is imitating an infinitesimally close individual, whereas in the two-type case, the imitation is at a distance. If we let w_θ be arbitrarily close to w_θ , then we would get e_θ and c_θ , as in the limit case of I .

Combining all of the effects gives

$$\begin{aligned} W'(x) &= -\xi\mu Y_\varphi(x)E_\varphi/x + \lambda(x)\hat{I} \\ &= -\frac{1}{x\sigma(x)} [\xi\mu Y_\varphi(x)E_\varphi(x)\sigma(x) + \hat{I}]. \end{aligned}$$

4.3 Marginal Tax Rates

At an optimum, $W'(x) = 0$, so

$$\xi = -\frac{\hat{I}/\sigma(x)}{Y_\varphi(x)E_\varphi}, \quad (30)$$

which coincides with the general formula (21) if $S = C = R = 0$ and when we replace I with $\hat{I}$. In particular, the fact that the discrete type model generically rules out two individuals with the same wage in different sectors implies $C = R = 0$. The exogeneity of

sectoral allocations eliminates the sectoral shift effect, so $S = 0$. Moreover, note that the sign of $\hat{l}$ is opposite of the sign of η . This means that $\xi > 0$ (and hence top marginal taxes are negative) precisely when we want to redistribute from the θ - to the φ -types, so that the downward incentive constraint binds.

This is analogous to our results in the general Roy model, but the addition of S and C will make ξ , and hence top marginal taxes, smaller in absolute value. To understand the intuition behind this, suppose we lower taxes at the top to increase the effort of the top earners. This is welfare enhancing because it raises the wages of the low wage sector workers and lowers the wages of high wage sector workers and thus relaxes the downward incentive constraint. However, when there is endogenous occupational choice, the sectoral shift effect works against this, since this change in wages leads some individuals to shift out of the high wage into the low wage sector, undoing some of the original increase in aggregate effort in the high wage sector. The indirect wage shift effect C is similar: the increase in wages in the low wage sector leads those who *do not shift* sectors to work harder, and the decrease in wages in the high wage sector leads those (again, who *do not shift* sectors) to work less hard. This also partially offsets the beneficial impact of increased effort among high earners.

These additional sectoral shift effects make the regressivity of the tax schedule that is optimal in the Stiglitz (1982) model less effective. As a result, optimal taxes in the general Roy model with continuous types, while less progressive than in a standard Mirrlees model, are more progressive than in a discrete type model without occupational choice.

5 A Numerical Example

In this section, we briefly illustrate our results with a simple numerical example. We assume quasilinear preferences $U(c, e) = c - h(e)$ with an isoelastic disutility $h(e) = e^{1+1/\varepsilon}/(1 + 1/\varepsilon)$ and an elasticity of labor supply $\varepsilon = 0.5$. The skill distribution has support $[\underline{\theta}, \bar{\theta}] \times [\underline{\varphi}, \bar{\varphi}] = [1, 16] \times [1, 6]$ and is independent across the two dimensions, so that $F(\theta, \varphi) = F_\theta(\theta)F_\varphi(\varphi)$. We assume that both F_θ and F_φ are Pareto distributions with parameters $\kappa_\theta = 2$ and $\kappa_\varphi = 8$. As a result, there is more mass on lower skills in the φ -dimension compared to θ , and we take Φ as the low skill and Θ as the high skill sector in the following. We truncate both distributions at the top and renormalize accordingly. Moreover, to prevent a kink in the wage distribution at $\bar{\varphi} = 6$, we shift f_φ so that $f_\varphi(\bar{\varphi}) = 0$ and renormalize accordingly.

Technology is given by a CES production function as in (13), so that $\sigma = 1/(1 - \rho)$ is the constant substitution elasticity. In particular, we start with the case $\rho = 0$ so that

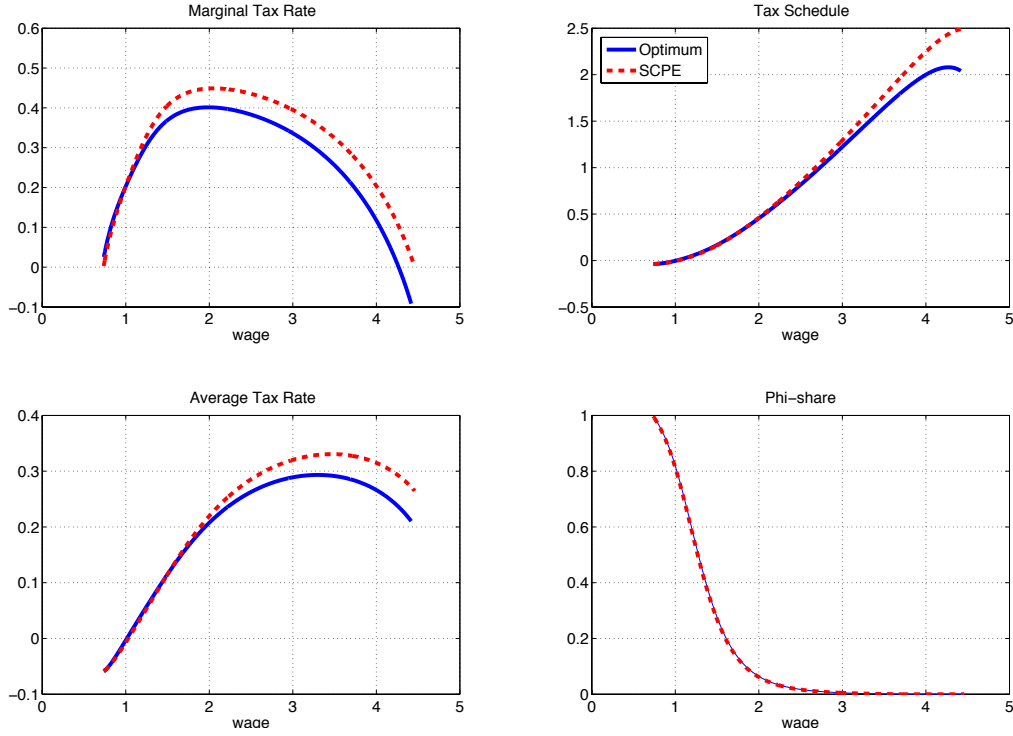


Figure 2: Marginal and average tax rates as a function of the wage

the substitution elasticity is one and technology is Cobb-Douglas. We set the aggregate income share of the high skill sector Θ to $\alpha = .2$.

Finally, we assume Pareto weights of the form $\Psi(F) = 1 - (1 - F)^r$. The parameter r thus characterizes the magnitude of the government's desire for redistribution from high to low wages: With quasilinear preferences, $r = 1$ implies no redistributive motives, and $r \rightarrow \infty$ for a Rawlsian social planner. We take $r = 1.3$, so that there is some intermediate desire for redistribution.

Figure 2 shows the marginal tax schedule $T'(y(w))$, the tax schedule $T(y(w))$, the average tax rate $T(y(w))/y(w)$ and the share of Φ -sector workers $f_x^\Phi(w)/f_x(w)$ as a function of the wage w both for the Pareto optimum and the SCPE resulting from our parametrization. It illustrates the patterns of the optimal tax policy derived so far: the top marginal tax rate is negative (-9.2%) and the optimal tax schedule is overall less progressive than in the SCPE. Figure 3 shows the same pattern as a function of income. Since the optimal tax schedule involves lower marginal tax rates at the top, it induces higher effort and hence the optimal income distribution extends to higher incomes than under the SCPE policy.

Figure 4 demonstrates that our assumptions are satisfied in the numerical example: individual effort is increasing in the wage w , and the share of Φ -sector workers is mono-

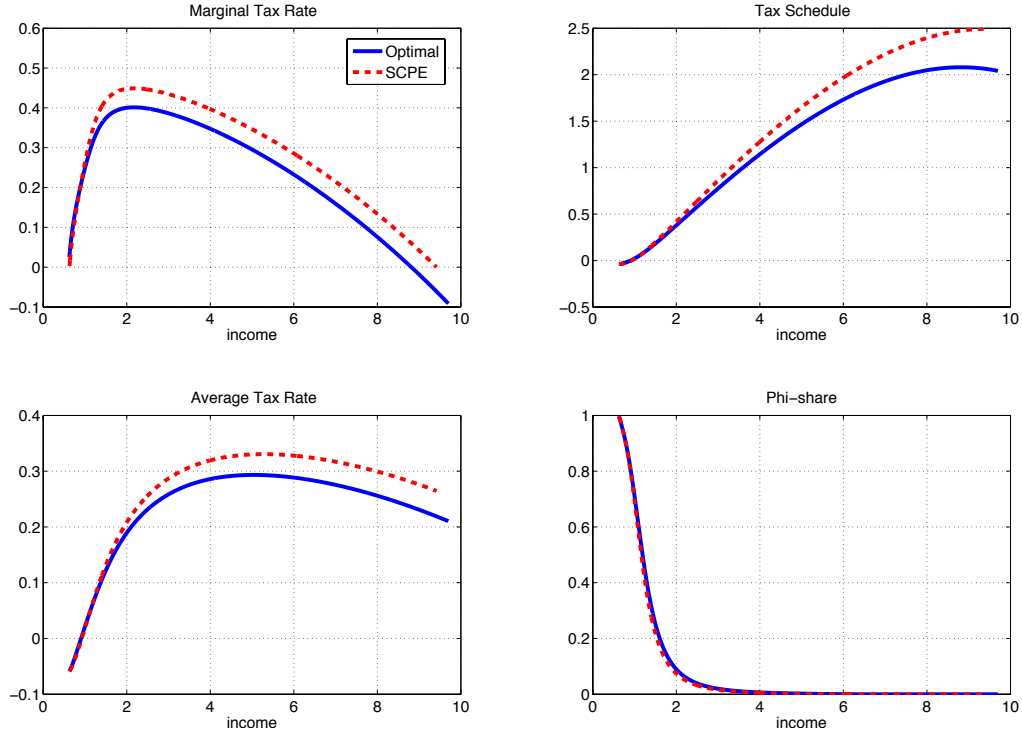


Figure 3: Marginal and average tax rates as a function of income

tone. A fortiori, income $y(w) = we(w)$ is increasing in the wage, so that bunching does not need to be considered.

In figure 5, we plot the marginal tax schedule for varying substitution elasticities $\sigma \in \{2, 1, 0.5, 0.1\}$, as a function of quantiles of the wage distribution. It shows that the tax schedule becomes less and less progressive as we move to lower substitution elasticities. This is because the general equilibrium effects from the endogeneity of wages become more pronounced as we move away from linear technology, with $\sigma = \infty$ and $\rho = 1$. As a result, the corresponding top marginal tax rates become more negative, moving from -7.6% to -9.2%, -10.8% and -11.9%.

6 Extensions

We consider two extensions here: allowing individuals to face different costs of working in different sectors, and an extension to unbounded skill distributions.

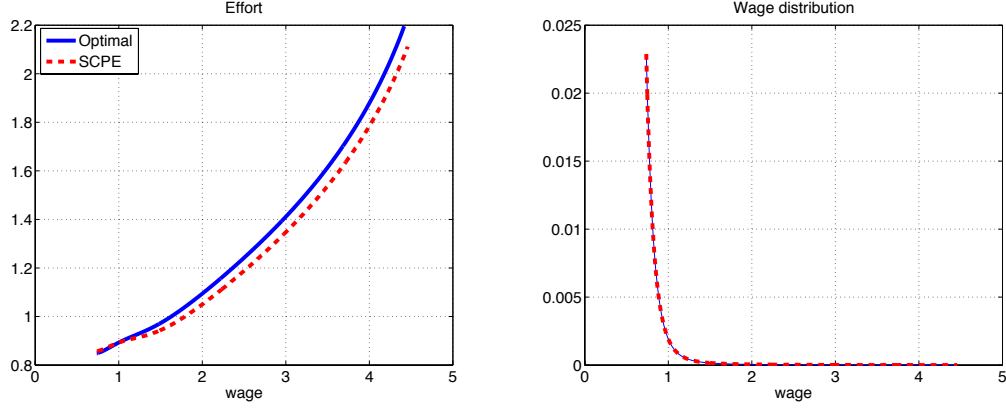


Figure 4: Effort and wage distribution

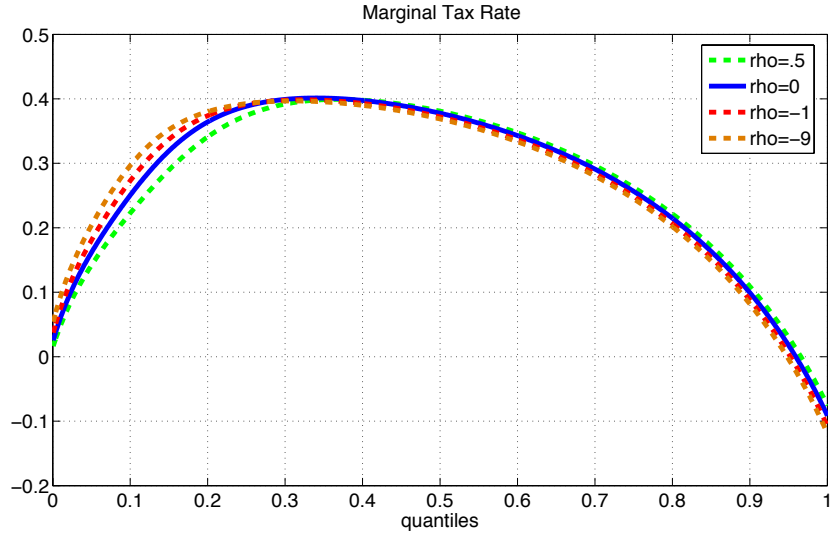


Figure 5: Marginal tax rates for varying σ as a function of wage quantiles

6.1 Differential Sectoral Costs

In the preceding analysis, individuals based their sectoral choice exclusively on whether the Θ - or Φ - sector afforded them a higher wage. It is straightforward to extend the model to an application in which individuals have different tastes for effort in the two sectors.

Let the types be described by a three-dimensional vector $t = (\theta, \varphi, \beta)$, where β is interpreted as an idiosyncratic cost of Θ -sector effort relative to Φ -sector effort and, as above, θ and φ measure the Θ - and Φ -sector skills. The type distribution has a general cdf $\hat{F}(\theta, \varphi, \beta)$, with support $[\underline{\theta}, \bar{\theta}] \times [\underline{\varphi}, \bar{\varphi}] \times [\underline{\beta}, \bar{\beta}]$, and cumulative welfare weights are given by $\hat{G}(\theta, \varphi, \beta)$.

Preferences are separable in effort and consumption, and isoelastic in effort:

$$U(c, e; t, P) = u(c) - \frac{e^{1+\varepsilon}}{1+1/\varepsilon} \times (\beta \mathbb{1}_P(t)), \quad (31)$$

where $\varepsilon = \varepsilon^u = \varepsilon^c$ is the elasticity of effort and P is the set of types working in the Θ -sector.

Although there are two dimensions of heterogeneity— θ and β —within the Θ -sector, as in Choné and Laroque (2010), these two dimensions collapse into a single-dimensional sufficient statistic. In particular, any two individuals with the same $\tilde{\theta} \equiv \theta / \beta^{\frac{\varepsilon}{1+\varepsilon}}$ have the same preferences over (c, y) -pairs and thus are observationally indistinguishable if they both choose to work in the Θ -sector. Moreover, any two types $(\theta_1, \varphi_1, \beta_1)$ and $(\theta_2, \varphi_2, \beta_2)$ with the same $\tilde{\theta}_1 = \tilde{\theta}_2$ and $\varphi_1 = \varphi_2$ make the same sectoral choice.

This means that we can “collapse” the policy relevant type distribution into the two-dimensional distribution of $(\tilde{\theta}, \varphi)$ types, with cdf

$$F(\tilde{\theta}, \varphi) \equiv \int_{\underline{\varphi}}^{\overline{\varphi}} \int_{\underline{\theta}}^{\theta} \int_{(\theta/\tilde{\theta})^{\frac{1+\varepsilon}{\varepsilon}}}^{\overline{\beta}} d\hat{F}(\theta', \varphi', \beta').$$

We can define the two-dimensional welfare weights $G(\tilde{\theta}, \varphi)$ analogously. By interpreting w as an *effective wage*, $\tilde{\theta}Y_\theta(x)$ or $\varphi Y_\varphi(x)$, the Pareto Optimal (and SCPE) tax rates are characterized exactly as in Section 3. In particular, marginal taxes are given by Proposition 1, with ξ as in equation (21), and, as in Proposition 2, top marginal tax rates will be negative whenever (a) the Θ -sector is the high wage sector, (b) effort is increasing in wage, (c) marginal social utility of consumption is decreasing in wage, and (d) there is a weak preference for Φ -sector workers at any given effective wage.

6.2 Asymptotics

We focused attention on bounded skill distributions for ease of interpretation and for comparison with Stiglitz (1982), but none of the analysis relies on this assumption. The analysis of the outer problem in section 3.3 can be extended in a straightforward way to unbounded distributions, and the methods developed in Saez (2001) can be used for asymptotics to extend Proposition 1 with general preferences. The asymptotics for the inner problem are particularly transparent in the special case considered in the following proposition.

Proposition 3. *Consider any Pareto optimum (respectively, SCPE) such that*

- (i) *preferences are quasilinear and isoelastic: $U(c, e) = c - \frac{e^{1+1/\varepsilon}}{1+1/\varepsilon}$*

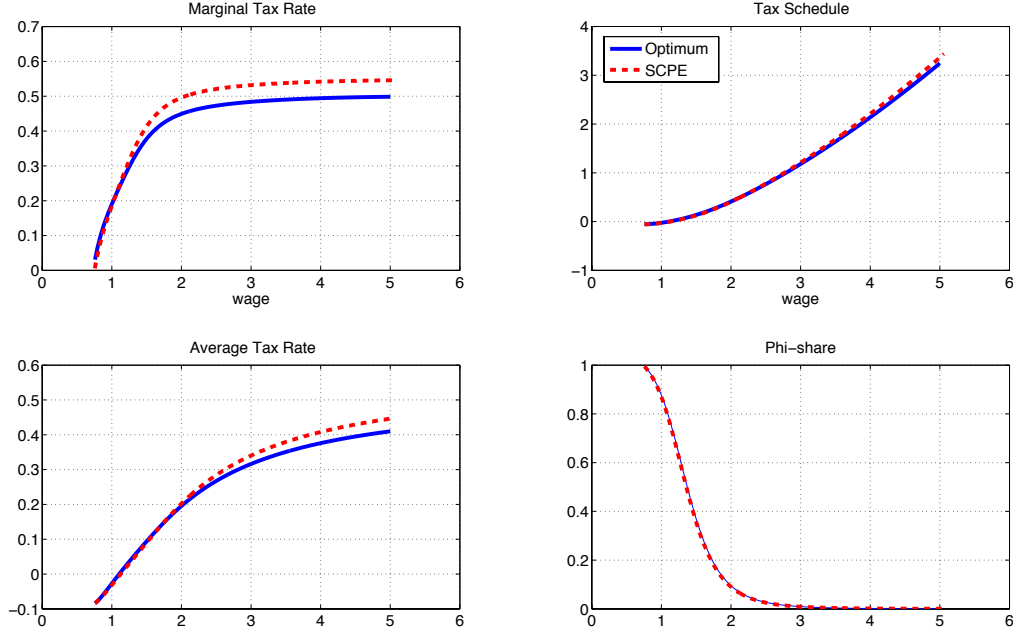


Figure 6: Marginal and average tax rates as a function of the wage

(ii) the top earners are all in the Θ -sector: $\lim_{w \rightarrow \infty} f_x^\Theta(w) / f_x(w) = 0$

(iii) the Θ -sector skill distribution has Pareto tails with parameter κ : $\lim_{w \rightarrow \infty} \frac{1 - F_x(w)}{w f_x(w)} = \kappa$

(iv) Pareto weights are relative and progressive: $G_x(w) = \Psi(F_x(w))$ with $\Psi''(x) < 0$ and

(vi) zero welfare weight is put on the top earners: $\Psi'(1) = 0$.

Then the top marginal tax rate is

$$\frac{\kappa (1 + 1/\varepsilon) - \xi \Gamma(x)}{\kappa (1 + 1/\varepsilon) + 1} \quad \left(\text{respectively, } \frac{\kappa (1 + 1/\varepsilon)}{\kappa (1 + 1/\varepsilon) + 1} \right), \quad (32)$$

and $\xi > 0$.

Proof. In Appendix B.5 □

In particular, this implies that asymptotic taxes are scaled down in the Pareto optimum relative to the SCPE. Figure 6 illustrates this for the baseline calibration from section 5.

7 Conclusion

We view this paper as making a two-fold contribution. The first is methodological: we provide a technique for solving a multi-dimensional screening problem in an important

class of contexts. Specifically, we show that the multi-dimensional screening problem that arises in designing optimal taxation in a multiple-sector economy can be reduced to a single dimensional optimal income tax problem à la Mirrlees. This basic technique is likely to be applicable more broadly.

Our second contribution is to derive some of the implications that self-selection into occupational sectors can have for optimal income taxation. In particular, we show that the presence of several complementary sectors in an economy provides a force pushing towards less progressive taxation.

Central to this result is our assumption—we believe quite reasonable—that the government cannot or chooses not to observe the sector an individual is employed in when it levies taxes. In ongoing research, we aim at relaxing this assumption and characterizing optimal sector-specific taxation.

References

- ALLEN, F. (1982): “Optimal Linear Income Taxation with General Equilibrium Effects on Wages,” *Journal of Public Economics*, 17, 135–143.
- BOADWAY, R., N. MARCEAU, AND P. PESTIEAU (1991): “Optimal Linear Income Taxation in Models with Occupational Choice,” *Journal of Public Economics*, 46, 133–162.
- BORJAS, G. (1987): “Self-Selection and the Earnings of Immigrants,” *American Economic Review*, 77, 531–553.
- (2002): “The Wage Structure and the Sorting of Workers into the Public Sector,” *NBER Working Paper* 9313.
- CHONÉ, P., AND G. LAROQUE (2010): “Negative Marginal Tax Rates and Heterogeneity,” *American Economic Review*, 100, 2532–2547.
- DIAMOND, P. (1998): “Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Tax Rates,” *American Economic Review*, 88, 83–95.
- FELDSTEIN, M. (1973): “On the Optimal Progressivity of the Income Tax,” *Journal of Public Economics*, 2, 356–376.
- FUDENBERG, D., AND D. LEVINE (2009): “Self-confirming equilibrium and the Lucas critique,” *Journal of Economic Theory*, 144, 2354–2371.

- KLEVEN, H. J., C. T. KREINER, AND E. SAEZ (2009): "The Optimal Income Taxation of Couples," *Econometrica*, 77, 537–560.
- LAROQUE, G. (2005): "Income maintenance and labor force participation," *Econometrica*, 73, 341–376.
- MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.
- MORESI, S. (1997): "Optimal Taxation and Firm Formation: A Model of Asymmetric Information," *European Economic Review*, 42, 1525–1551.
- NAITO, H. (1999): "Re-Examination of Uniform Commodity Taxes Under A Non-Linear Income Tax System And Its Implication for Production Efficiency," *Journal of Public Economics*, 71, 165–188.
- PARKER, S. C. (1999): "The Optimal Linear Taxation of Employment and Self-Employment Incomes," *Journal of Public Economics*, 73, 107–123.
- ROCHET, J.-C., AND P. CHONÉ (1998): "Ironing, Sweeping and Multidimensional Screening," *Econometrica*, 66, 783–826.
- ROTHSCHILD, C., AND F. SCHEUER (2011): "Optimal Taxation with Rent-Seeking," *NBER Working Paper 17035*.
- ROY, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.
- SAEZ, E. (2001): "Using Elasticities to Derive Optimal Tax Rates," *Review of Economic Studies*, 68, 205–229.
- SARGENT, T. (2008): "Evolution and Intelligent Design," *American Economic Review*, 98, 5–37.
- STIGLITZ, J. (1982): "Self-Selection and Pareto Efficient Taxation," *Journal of Public Economics*, 17, 213–240.
- WERNING, I. (2007): "Pareto-Efficient Income Taxation," *Mimeo, MIT*.
- ZECKHAUSER, R. (1977): "Taxes in Fantasy, or Most Any Tax on Labor Can Turn Out to Help the Laborers," *Journal of Public Economics*, 8, 133–150.

A Proofs for Section 2

A.1 Proof of Lemma 1

We prove the result in the following four steps:

Step 1. It is an immediate consequence of the incentive constraints that the utility of a type sent to a given sector can only depend on his wage in that sector. Formally, consider two types (θ_0, φ_0) and (θ_1, φ_1) with

$$S(\theta_0, \varphi_0) = S(\theta_1, \varphi_1) = \Theta \text{ and } Y_\theta(x)\theta_0 = Y_\theta(x)\theta_1 = w.$$

Whenever

$$U\left(c(w/Y_\theta(x), \varphi_0), \frac{y(w/Y_\theta(x), \varphi_0)}{w}\right) \neq U\left(c(w/Y_\theta(x), \varphi_1), \frac{y(w/Y_\theta(x), \varphi_1)}{w}\right),$$

either type $(w/Y_\theta(x), \varphi_0)$'s or type $(w/Y_\theta(x), \varphi_1)$'s incentive constraint is violated. An analogous argument applies to types sent to the Φ -sector. Since w is the wage of the agent in the sector that he is sent to, we can thus write utilities as $V_\theta(w)$ for all types for which $S(\theta, \varphi) = \Theta$ and $V_\varphi(w)$ for all types with $S(\theta, \varphi) = \Phi$.

Step 2. We now show that the consumption and income allocated to a type who is sent to a given sector can also only depend on his skill in that sector. To see this, consider again two types (θ_0, φ_0) and (θ_1, φ_1) with

$$S(\theta_0, \varphi_0) = S(\theta_1, \varphi_1) = \Theta \text{ and } \theta_0 = \theta_1 = w/Y_\theta(x).$$

Consider the expression

$$\begin{aligned} H(w, w') &= \left[U\left(c(w'/Y_\theta(x), \varphi_0), \frac{y(w'/Y_\theta(x), \varphi_0)}{w'}\right) - U\left(c(w/Y_\theta(x), \varphi_0), \frac{y(w/Y_\theta(x), \varphi_0)}{w'}\right) \right] \\ &\quad - \left[U\left(c(w'/Y_\theta(x), \varphi_1), \frac{y(w'/Y_\theta(x), \varphi_1)}{w'}\right) - U\left(c(w/Y_\theta(x), \varphi_1), \frac{y(w/Y_\theta(x), \varphi_1)}{w'}\right) \right]. \end{aligned}$$

Mechanically, we have

$$\left. \frac{\partial H(w, w')}{\partial w'} \right|_{w'=w} = 0,$$

since the partial derivative of each bracketed term, evaluated at $w' = w$, is individually zero. On the other hand, we showed in 1. that

$$U\left(c(w'/Y_\theta(x), \varphi_0), \frac{y(w'/Y_\theta(x), \varphi_0)}{w'}\right) = U\left(c(w'/Y_\theta(x), \varphi_1), \frac{y(w'/Y_\theta(x), \varphi_1)}{w'}\right) = V_\theta(w'),$$

so that $H(w, w')$ reduces to

$$H(w, w') = U\left(c(w/Y_\theta(x), \varphi_1), \frac{y(w/Y_\theta(x), \varphi_1)}{w'}\right) - U\left(c(w/Y_\theta(x), \varphi_0), \frac{y(w/Y_\theta(x), \varphi_0)}{w'}\right).$$

Single-crossing implies that

$$U_w\left(c(w/Y_\theta(x), \varphi_1), \frac{y(w/Y_\theta(x), \varphi_1)}{w}\right) \geq U_w\left(c(w/Y_\theta(x), \varphi_0), \frac{y(w/Y_\theta(x), \varphi_0)}{w}\right)$$

whenever

$$U\left(c(w/Y_\theta(x), \varphi_1), \frac{y(w/Y_\theta(x), \varphi_1)}{w}\right) = U\left(c(w/Y_\theta(x), \varphi_0), \frac{y(w/Y_\theta(x), \varphi_0)}{w}\right)$$

and $c(w/Y_\theta(x), \varphi_1) \geq c(w/Y_\theta(x), \varphi_0)$.¹⁰ Hence, it must be that

$$c(w/Y_\theta(x), \varphi_1) = c(w/Y_\theta(x), \varphi_0) \text{ and } y(w/Y_\theta(x), \varphi_1) = y(w/Y_\theta(x), \varphi_0).$$

The same argument applies to types sent to the Φ -sector. We can thus write allocations as $c_\theta(w), y_\theta(w)$ for all types with $S(\theta, \varphi) = \Theta$ and $c_\varphi(w), y_\varphi(w)$ for all types with $S(\theta, \varphi) = \Phi$.

Step 3. It is also straightforward to see that two types who earn the same wage in different sectors must get the same utility. Consider two types (θ_0, φ_0) and (θ_1, φ_1) with

$$S(\theta_0, \varphi_0) = \Theta, S(\theta_1, \varphi_1) = \Phi \text{ and } Y_\theta(x)\theta_0 = Y_\varphi(x)\varphi_1 = w.$$

Next, assume w.l.o.g.

$$U\left(c_\theta(w), \frac{y_\theta(w)}{w}\right) > U\left(c_\varphi(w), \frac{y_\varphi(w)}{w}\right).$$

Then the $(\theta_1, w/Y_\varphi(x))$ -type could imitate the $(w/Y_\theta(x), \varphi_0)$ -type by producing income $y_\theta(w)$ but doing so in the Φ -sector, i.e. using his skill $w/Y_\varphi(x)$. His utility from this deviation would be exactly the LHS of the above inequality, contradicting incentive compatibility. This shows

$$V_\theta(w) = V_\varphi(w) \equiv V(w) \forall w.$$

Step 4. We finally show that the incentive constraints also imply that allocations have to be such that

$$c_\theta(w) = c_\varphi(w) \equiv c(w) \text{ and } y_\theta(w) = y_\varphi(w) \equiv y(w) \forall w,$$

i.e. two types who earn the same wage in different sectors must get the same consumption and income. For this purpose, consider again the expression

$$\begin{aligned} H(w, w') &= \left[U\left(c_\theta(w'), \frac{y_\theta(w')}{w'}\right) - U\left(c_\theta(w), \frac{y_\theta(w)}{w}\right) \right] \\ &\quad - \left[U\left(c_\varphi(w'), \frac{y_\varphi(w')}{w'}\right) - U\left(c_\varphi(w), \frac{y_\varphi(w)}{w}\right) \right]. \end{aligned}$$

Now consider again $\partial H / \partial w' |_{w'=w}$. On the one hand, this is mechanically zero since the partial derivative of each bracketed term is individually zero. On the other hand,

$$U\left(c_\theta(w'), \frac{y_\theta(w')}{w'}\right) = U\left(c_\varphi(w'), \frac{y_\varphi(w')}{w'}\right)$$

¹⁰Recall that single crossing implies that $-U_e/(wU_c)$ is decreasing in w , or equivalently that U_y/U_c is increasing in w and hence $U_{yw}U_c - U_yU_{cw} > 0$. Therefore, the change in U_w from a marginal increase in c and y along w 's indifference curve (i.e. such that $dc = -(U_y/U_c)dy$) is $dU_w = U_{wy} - (U_y/U_c)U_{wc} > 0$.

by step 3., so

$$H(w, w') = U\left(c_\varphi(w), \frac{y_\varphi(w)}{w'}\right) - U\left(c_\theta(w), \frac{y_\theta(w)}{w'}\right),$$

and, by single crossing

$$U_w\left(c_\varphi(w), \frac{y_\varphi(w)}{w}\right) \geq U_w\left(c_\theta(w), \frac{y_\theta(w)}{w}\right)$$

whenever $U(c_\varphi(w), y_\varphi(w)/w) = U(c_\theta(w), y_\theta(w)/w)$ and $c_\varphi(w) \geq c_\theta(w)$. Hence, it must be that $c_\varphi(w) = c_\theta(w)$ and $y_\varphi(w) = y_\theta(w)$ for all w .

These four steps prove that any incentive compatible direct mechanism in our economy can be implemented by offering a schedule of $c(w), y(w)$ bundles with $w \equiv \max\{Y_\theta(x)\theta, Y_\varphi(x)\varphi\}$, which is equivalent to a single non-linear tax schedule.

B Proofs for Section 3

B.1 Proof of Lemma 2

By changing variables to $p \equiv F_x(w)$, we can write the planner's welfare in any feasible allocation as

$$SW = \int_{\underline{w}_x}^{\bar{w}_x} \psi(F_x(w)) V(w) dF_x(w) = \int_0^1 \psi(p) V(F_x^{-1}(p)) dp$$

with $\psi(F) = \Psi'(F)$. Note that the welfare weights $\psi(p)$ are independent of the allocation and that SW depends only on the *distribution* of V in this formulation. It follows that if the distribution of V under one feasible allocation first order stochastically dominates (FOSD) the distribution of V under a second feasible allocation, then social welfare is higher under the first allocation. To complete the proof, note that $\{x_0, e_0(\theta, \varphi), V_0(\theta, \varphi)\}$ cannot maximize social welfare if it is Pareto dominated by a feasible allocation $\{x_1, e_1(\theta, \varphi), V_1(\theta, \varphi)\}$, as the distribution of V_1 would then FOSD the distribution of V_0 .

B.2 Proof of Proposition 1

Putting multipliers μ on (6), $\zeta\lambda$ on (5) and $\hat{\eta}(w)\mu$ on (4), the Lagrangian corresponding to (3)-(6) is, after integrating by parts (4),

$$\begin{aligned} \mathcal{L} = & \int_{\underline{w}_x}^{\bar{w}_x} V(w) g_x(w) dw - \int_{\underline{w}_x}^{\bar{w}_x} V(w) \hat{\eta}'(w) \mu dw + \int_{\underline{w}_x}^{\bar{w}_x} U_e(c(V(w), e(w)), e(w)) \frac{e(w)}{w} \hat{\eta}(w) \mu dw \\ & + \zeta \mu \Gamma(x) \int_{\underline{w}_x}^{\bar{w}_x} w e(w) f_x^\theta(w) dw - \zeta \mu \int_{\underline{w}_x}^{\bar{w}_x} w e(w) f_x^\varphi(w) dw \\ & + \mu \int_{\underline{w}_x}^{\bar{w}_x} w e(w) f_x(w) dw - \mu \int_{\underline{w}_x}^{\bar{w}_x} c(V(w), e(w)) f_x(w) dw. \end{aligned} \quad (33)$$

Using $\partial c / \partial V = 1/U_c$ and compressing notation, the first order condition for $V(w)$ is

$$\hat{\eta}'(w) \mu = g_x(w) - \mu f_x(w) \frac{1}{U_c(w)} + \hat{\eta}(w) \mu \frac{U_{ec}(w)}{U_c(w)} \frac{e(w)}{w}. \quad (34)$$

Defining $\eta(w) \equiv \hat{\eta}(w)U_c(w)$, this becomes

$$\eta'(w) = g_x(w) \frac{U_c(w)}{\mu} - f_x(w) + \eta(w) \frac{U_{cc}(w)c'(w) + U_{ce}(w)e'(w) + U_{ce}(w)e(w)/w}{U_c(w)}. \quad (35)$$

Using the first order condition corresponding to the optimization problem for an individual worker,

$$U_c(w)c'(w) + U_e(w)e'(w) + U_e(w) \frac{e(w)}{w} = 0,$$

the fraction in (35) can be written as $-(\partial MRS(w)/\partial c)y'(w)/w$ where

$$MRS(w) \equiv -\frac{U_e(c(w), e(w))}{U_c(c(w), e(w))}$$

is the marginal rate of substitution between effort and consumption. Substituting in (35) and rearranging yields

$$-\frac{\partial MRS(w)}{\partial c} e(w) \frac{y'(w)}{y(w)} \eta(w) = f_E(w) - \psi_E(w) \frac{U_c(w)}{\mu} + \eta'(w). \quad (36)$$

Integrating this ODE gives

$$\begin{aligned} \eta(w) &= \int_w^{\bar{w}_x} \left(f_x(z) - \psi_x(z) \frac{U_c(z)}{\mu} \right) \exp \left(\int_w^z \frac{\partial MRS(s)}{\partial c} e(s) \frac{y'(s)}{y(s)} ds \right) dz \\ &= \int_w^{\bar{w}_x} \left(1 - \frac{\psi_x(z)}{f_x(z)} \frac{U_c(z)}{\mu} \right) \exp \left(\int_w^z \left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)} \right) \frac{dy(s)}{y(s)} \right) f_x(z) dz, \end{aligned} \quad (37)$$

where the last step follows from $e(w)\partial MRS(w)/\partial c = 1 - \varepsilon^u(w)/\varepsilon^c(w)$ after tedious algebra (e.g. using equations (23) and (24) in Saez, 2001).

Using $\partial c/\partial e = MRS$, the first order condition for $e(w)$ is

$$\begin{aligned} &\mu w f_x(w) \left(1 - \frac{MRS(w)}{w} \right) - \xi \mu w \left(\Gamma(x) f_x^\theta(w) - f_x^\varphi(w) \right) \\ &= -\hat{\eta}(w) \mu \left[\frac{(-U_{ec}(w)U_e(w)/U_c(w) + U_{ee}(w))e(w)}{w} + \frac{U_e(w)}{w} \right], \end{aligned}$$

which after some algebra can be rewritten as

$$w f_x(w) \left(1 - \frac{MRS(w)}{w} \right) + \xi w \left(\Gamma(x) f_x^\theta(w) - f_x^\varphi(w) \right) = \eta(w) \left(\frac{\partial MRS(w)}{\partial e} \frac{e}{w} + \frac{MRS(w)}{w} \right). \quad (38)$$

Noting that $MRS(w)/w = 1 - T'(y(w))$ from the first order condition of the workers' utility maximization problem and using the definition of $\eta(w)$, this becomes

$$1 + \xi \frac{\Gamma(x) f_x^\theta(w) - f_x^\varphi(w)}{f_x(w)} = (1 - T'(y(w))) \left[1 + \frac{\eta(w)}{w f_x(w)} \left(1 + \frac{\partial MRS(w)}{\partial e} \frac{e}{MRS(w)} \right) \right]. \quad (39)$$

Simple algebra again shows that $1 + \partial \log MRS(w)/\partial \log e = (1 + \varepsilon^u(w))/\varepsilon^c(w)$, and that

$$\frac{\Gamma(x) f_x^\theta(w) - f_x^\varphi(w)}{f_x(w)} = \frac{Y_\varphi(x)}{Y_\varphi(x) + x Y_\theta(x)} - \frac{f_x^\varphi(w)}{f_x(w)}.$$

The Proposition follows from (37) and (39).

B.3 Proof of Lemma 3

The substitution elasticity is defined as

$$\begin{aligned}\sigma(x) &\equiv \frac{dx}{x} \frac{Y_\varphi(x)/Y_\theta(x)}{d(Y_\varphi(x)/Y_\theta(x))} = \frac{1}{\frac{d(Y_\varphi(x)/Y_\theta(x))}{dx}} \frac{Y_\varphi(x)}{xY_\theta(x)} \\ &= \frac{Y_\theta(x)Y_\varphi(x)}{x} \frac{1}{Y_\theta(x)Y'_\varphi(x) - Y_\varphi(x)Y'_\theta(x)} = \frac{1}{x \frac{Y'_\varphi(x)}{Y_\varphi(x)} - x \frac{Y'_\theta(x)}{Y_\theta(x)}} = -\frac{1}{x\lambda(x)}.\end{aligned}$$

B.4 Proof of Lemma 4

We first state the following technical lemma—the proof of which involves nothing more than tedious algebra.

Lemma 5.

$$\frac{dF_x^\theta(w)}{dx} = -\frac{Y'_\theta(x)}{Y_\theta(x)} w f_x^\theta(w) + \Omega_x(w) \quad \text{and} \quad \frac{dF_x^\varphi(w)}{dx} = -\frac{Y'_\varphi(x)}{Y_\varphi(x)} w f_x^\varphi(w) - \Omega_x(w)$$

with

$$\Omega_x(w) \equiv \frac{1}{Y_\theta(x)Y_\varphi(x)} \lambda(x) \int_{\underline{w}_x}^w w' f\left(\frac{w'}{Y_\theta(x)}, \frac{w'}{Y_\varphi(x)}\right) dw'.$$

Completely analogous expressions hold for $G_x^\theta(w)$ and $G_x^\varphi(w)$.

This will be useful in the proof of Lemma 4, to which we now turn. Using (33),

$$\begin{aligned}W'(x) &= \int_{\underline{w}_x}^{\bar{w}_x} V(w) \frac{dG_x(w)}{dx} dw - \mu \int_{\underline{w}_x}^{\bar{w}_x} c(V(w), e(w)) \frac{dF_x(w)}{dx} dw + \xi \mu \Gamma'(x) Y_\theta(x) E_\theta \\ &\quad + \mu \xi \left(\Gamma(x) \int_{\underline{w}_x}^{\bar{w}_x} w e(w) \frac{dF_x^\theta(w)}{dx} dw - \int_{\underline{w}_x}^{\bar{w}_x} w e(w) \frac{dF_x^\varphi(w)}{dx} dw \right) + \mu \int_{\underline{w}_x}^{\bar{w}_x} w e(w) \frac{dF_x(w)}{dx} dw + B_1\end{aligned}$$

with

$$\begin{aligned}B_1 &\equiv \frac{d\bar{w}_x}{dx} \left[V(\bar{w}_x) g_x(\bar{w}_x) - \mu c(V(\bar{w}_x), e(\bar{w}_x)) f_x(\bar{w}_x) + \mu \left(f_x(\bar{w}_x) + \xi \left(\Gamma(x) f_x^\theta(e(\bar{w}_x)) - f_x^\varphi(e(\bar{w}_x)) \right) \right) \bar{w}_x e(\bar{w}_x) \right] \\ &\quad - \frac{d\underline{w}_x}{dx} \left[V(\underline{w}_x) g_x(\underline{w}_x) - \mu c(V(\underline{w}_x), e(\underline{w}_x)) f_x(\underline{w}_x) + \mu \left(f_x(\underline{w}_x) + \xi \left(\Gamma(x) f_x^\theta(e(\underline{w}_x)) - f_x^\varphi(e(\underline{w}_x)) \right) \right) \underline{w}_x e(\underline{w}_x) \right].\end{aligned}$$

Integrating by parts the five integral terms yields

$$\begin{aligned}W'(x) &= B_1 + B_2 - \int_{\underline{w}_x}^{\bar{w}_x} V'(w) \frac{dG_x(w)}{dx} dw + \mu \int_{\underline{w}_x}^{\bar{w}_x} \left(\frac{V'(w)}{U_c(w)} + MRS(w) e'(w) \right) \frac{dF_x(w)}{dx} dw + \xi \mu \Gamma'(x) Y_\theta(x) E_\theta \\ &\quad + \mu \xi \left(\int_{\underline{w}_x}^{\bar{w}_x} (w e'(w) + e(w)) \left(\frac{dF_x^\varphi(w)}{dx} - \Gamma(x) \frac{dF_x^\theta(w)}{dx} \right) dw \right) - \mu \int_{\underline{w}_x}^{\bar{w}_x} (w e'(w) + e(w)) \frac{dF_x(w)}{dx} dw\end{aligned}\tag{40}$$

with

$$B_2 = \left[V(w) \frac{dG_x(w)}{dx} - \mu c(V(w), e(w)) \frac{dF_x(w)}{dx} + \mu \xi w e(w) \left(\Gamma(x) \frac{dF_x^\theta(w)}{dx} - \frac{dF_x^\varphi(w)}{dx} \right) + \mu w e(w) \frac{dF_x(w)}{dx} \right]_{\underline{w}_x}^{\bar{w}_x}.$$

By the first order conditions (36) and (38) with respect to $V(w)$ and $e(w)$ from the inner problem, the terms

$$\mu \int_{\underline{w}_x}^{\bar{w}_x} \frac{e'(w)}{f_x(w)} \left[w f_x(w) \left(1 - \frac{MRS(w)}{w} \right) + \xi w \left(\Gamma(x) f_x^\theta - f_x^\varphi(w) \right) - \eta(w) \left(\frac{\partial MRS(w)}{\partial e} \frac{e(w)}{w} + \frac{MRS(w)}{w} \right) \right] \frac{dF_x(w)}{dx} dw$$

and

$$\mu \int_{\underline{w}_x}^{\bar{w}_x} \frac{V'(w)}{U_c(w) f_x(w)} \left[g_x(w) \frac{U_c(w)}{\lambda} - f_x(w) - \eta'(w) - \eta(w) \frac{\partial MRS(w)}{\partial c} e(w) \frac{y'(w)}{y(w)} \right] \frac{dF_x(w)}{dx} dw$$

are both equal to zero. Adding them to (40), using (4) and re-arranging yields

$$\begin{aligned} W'(x) = & B_1 + B_2 + \xi \mu \Gamma'(x) Y_\theta(x) E_\theta + \int_{\underline{w}_x}^{\bar{w}_x} V'(w) \left(\frac{g_x(w)}{f_x(w)} \frac{dF_x(w)}{dx} - \frac{dG_x(w)}{dx} \right) dw \\ & - \mu \int_{\underline{w}_x}^{\bar{w}_x} e(w) \frac{dF_x(w)}{dx} dw - \mu \int_{\underline{w}_x}^{\bar{w}_x} \left(\frac{\eta(w)}{w} \frac{d[MRS(w)e(w)]}{dw} + \eta'(w) \frac{V'(w)}{U_c(w)} \right) \frac{1}{f_x(w)} \frac{dF_x(w)}{dx} dw \\ & + \xi \mu \int_{\underline{w}_x}^{\bar{w}_x} \left((e(w) + w e'(w)) \left(\frac{dF_x^\varphi(w)}{dx} - \Gamma(x) \frac{dF_x^\theta(w)}{dx} \right) + w e'(w) \left(\Gamma(x) \frac{f_x^\theta(w)}{f_x(w)} - \frac{f_x^\varphi(w)}{f_x(w)} \right) \frac{dF_x(w)}{dx} \right) dw. \end{aligned} \quad (41)$$

From Lemma 5,

$$\begin{aligned} \frac{g_x(w)}{f_x(w)} \frac{dF_x(w)}{dx} - \frac{dG_x(w)}{dx} &= -\frac{Y'_\theta(x)}{Y_\theta(x)} \left[\frac{g_x(w)}{f_x(w)} f_x^\theta(w) - g_x^\theta(w) \right] + \frac{Y'_\varphi(x)}{Y_\varphi(x)} \left[\frac{g_x(w)}{f_x(w)} f_x^\varphi(w) - g_x^\varphi(w) \right] \\ &= \left(\frac{Y'_\theta(x)}{Y_\theta(x)} - \frac{Y'_\varphi(x)}{Y_\varphi(x)} \right) \left[\frac{g_x(w)}{f_x(w)} f_x^\varphi(w) - g_x^\varphi(w) \right] \\ &= \lambda(x) \frac{f_x^\varphi(w) f_x^\theta(w)}{f_x(w)} \left(\frac{g_x^\theta(w)}{f_x^\theta(w)} - \frac{g_x^\varphi(w)}{f_x^\varphi(w)} \right). \end{aligned}$$

The first integral in (41) is therefore

$$\int_{\underline{w}_x}^{\bar{w}_x} V'(w) \left(\frac{g_x(w)}{f_x(w)} \frac{dF_x(w)}{dx} - \frac{dG_x(w)}{dx} \right) dw = \lambda(x) R \quad (42)$$

Combining the terms with $e(w)$ on the second and third line of (41) and using lemma 5 and the identity

$Y'_\theta(x)E_\theta + Y'_\varphi(x)E_\varphi \equiv 0$ gives:

$$\begin{aligned}
& -\mu \int_{\underline{w}_x}^{\bar{w}_x} e(w) \frac{dF_x(w)}{dx} dw + \xi \mu \int_{\underline{w}_x}^{\bar{w}_x} e(w) \left(\frac{dF_x^\varphi(w)}{dx} - \Gamma(x) \frac{dF_x^\theta(w)}{dx} \right) dw \\
& = \mu \left[\frac{Y'_\theta(x)}{Y_\theta(x)} \int_{\underline{w}_x}^{\bar{w}_x} we(w) f_x^\theta(w) dw + \frac{Y'_\varphi(x)}{Y_\varphi(x)} \int_{\underline{w}_x}^{\bar{w}_x} we(w) f_x^\varphi(w) dw \right] \\
& \quad + \xi \mu \int_{\underline{w}_x}^{\bar{w}_x} we(w) \left(\Gamma(x) \frac{Y'_\theta(x)}{Y_\theta(x)} f_x^\theta(w) - \frac{Y'_\varphi(x)}{Y_\varphi(x)} f_x^\varphi(w) \right) dw - \xi \mu (1 + \Gamma(x)) \int_{\underline{w}_x}^{\bar{w}_x} e(w) \Omega_x(w) dw \\
& = 0 + \xi \mu (1 + \Gamma(x)) Y'_\theta E_\theta - \xi \mu (1 + \Gamma(x)) \int_{\underline{w}_x}^{\bar{w}_x} e(w) \Omega_x(w) dw. \tag{43}
\end{aligned}$$

The terms with $we'(w)$ in the last line (41) can be written, using Lemma 5 again, as

$$\begin{aligned}
& \xi \mu \int_{\underline{w}_x}^{\bar{w}_x} we'(w) \left(\frac{dF_x^\varphi(w)}{dx} - \Gamma(x) \frac{dF_x^\theta(w)}{dx} \right) + we'(w) \left(\Gamma(x) \frac{f_x^\theta(w)}{f_x(w)} - \frac{f_x^\varphi(w)}{f_x(w)} \right) \frac{dF_x(w)}{dx} \\
& = \xi \mu \int_{\underline{w}_x}^{\bar{w}_x} w^2 e'(w) \left[\Gamma(x) \frac{Y'_\theta(x)}{Y_\theta(x)} f_x^\theta(w) - \frac{Y'_\varphi(x)}{Y_\varphi(x)} f_x^\varphi(w) - \left(\Gamma(x) \frac{f_x^\theta(w)}{f_x(w)} - \frac{f_x^\varphi(w)}{f_x(w)} \right) \left(\frac{Y'_\theta(x)}{Y_\theta(x)} f_x^\theta(w) + \frac{Y'_\varphi(x)}{Y_\varphi(x)} f_x^\varphi(w) \right) \right] dw \\
& \quad - \xi \mu (1 + \Gamma(x)) \int_{\underline{w}_x}^{\bar{w}_x} we'(w) \Omega_x(w) dw \\
& = \lambda(x) \xi \mu (1 + \Gamma(x)) C - \xi \mu (1 + \Gamma(x)) \int_{\underline{w}_x}^{\bar{w}_x} we'(w) \Omega_x(w) dw, \tag{44}
\end{aligned}$$

where the first term in the last step follows after some tedious algebra.

Combining the terms with $\Omega_x(w)$ from (43) and (44) gives $-\xi \mu (1 + \Gamma(x)) \int_{\underline{w}_x}^{\bar{w}_x} (we(w))' \Omega_x(w) dw$, which can be integrated by parts to yield:

$$-B_3 + (1 + \Gamma(x)) \xi \mu \frac{\lambda(x)}{Y_\theta(x) Y_\varphi(x)} \int_{\underline{w}_x}^{\bar{w}_x} w^2 e(w) f \left(\frac{w}{Y_\theta(x)}, \frac{w}{Y_\varphi(x)} \right) dw = (1 + \Gamma(x)) \xi \mu \lambda(x) S, \tag{45}$$

with $B_3 = (1 + \Gamma(x)) \xi \mu \bar{w}_x e(\bar{w}_x) \Omega_x(\bar{w}_x)$ since $\Omega_x(\underline{w}_x) = 0$.

Finally, use the incentive constraint (4), rewritten as $V'(w)/U_c(w) = MRS(w)e(w)$, to write the last line of (41) as

$$-\lambda \int_{\underline{w}_x}^{\bar{w}_x} \left(\eta(w) w \frac{d[V'(w)/U_c(w)]}{dw} + \eta'(w) w \frac{V'(w)}{U_c(w)} + \eta(w) \frac{V'(w)}{U_c(w)} \right) \frac{1}{w f_x(w)} \frac{dF_x(w)}{dx} dw$$

or, recognizing the sum of the bracketed terms as $d[\eta(w)wV'(w)/U_c(w)]/dw$, integrating by parts, and using the transversality condition $\eta(\underline{w}_x) = \eta(\bar{w}_x) = 0$ and Lemma 5,

$$\mu \int_{\underline{w}_x}^{\bar{w}_x} \eta(w) w \frac{V'(w)}{U_c(w)} \frac{d}{dw} \left(- \left(\frac{Y'_\theta(x)}{Y_\theta(x)} \frac{f_x^\theta(w)}{f_x(w)} - \frac{Y'_\varphi(x)}{Y_\varphi(x)} \frac{f_x^\varphi(w)}{f_x(w)} \right) \right) dw = \mu \lambda(x) I \tag{46}$$

Define $\tilde{F}(w, x) \equiv F_x(w)$. Since $\tilde{F}(\bar{w}_x, x) \equiv 1$ for all x ,

$$\frac{d\tilde{F}(\bar{w}_x, x)}{dx} = \frac{\partial \tilde{F}(\bar{w}_x, x)}{\partial x} + \frac{\partial \tilde{F}(\bar{w}_x, x)}{\partial w} \frac{d\bar{w}_x}{dx} = \frac{dF_x(\bar{w}_x)}{dx} + f_x(\bar{w}_x) \frac{d\bar{w}_x}{dx} = 0. \tag{47}$$

Together with an analogous expression at $\underline{w}_x$, the fact that $\Omega_x(\underline{w}_x) = 0$, and Lemma 5, this yields

$$B_1 + B_2 = -\xi\mu\bar{w}_xe(\bar{w}_x)\Omega_x(\bar{w}_x) = B_3.$$

Using (42), (43), (44), (45) and (46) in (41) yields

$$W'(E) = \lambda(x) \left(I + R + \xi\mu \left[Y_\varphi(x)E_\varphi\sigma(x) + (1 + \Gamma(x))(S + C) \right] \right), \quad (48)$$

where we have used $(\Gamma'(x)Y_\theta(x) + (1 + \Gamma(x))Y'_\theta(x))E_\theta = -Y_\varphi(x)E_\varphi/x$.

B.5 Proof of Proposition 3

$$\begin{aligned} \lim_{w \rightarrow \infty} 1 - T'(y(w)) &= \lim_{w \rightarrow \infty} \frac{1 + \xi(1 + \Gamma(x)) \left(\frac{Y_\varphi(x)}{Y_\varphi(x) + xY_\theta(x)} - \frac{f_x^\varphi(w)}{f_x(w)} \right)}{1 + \frac{\eta(w)}{wf_x(w)} \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)}} \\ &= \frac{1 + \xi(1 + \Gamma(x)) \left(\frac{Y_\varphi(x)}{Y_\varphi(x) + xY_\theta(x)} - \lim_{w \rightarrow \infty} \left(\frac{f_x^\varphi(w)}{f_x(w)} \right) \right)}{1 + \left(1 + \frac{1}{\varepsilon} \right) \lim_{w \rightarrow \infty} \left(\frac{1 - F_x(w)}{wf_x(w)} \right) \left(\frac{\Psi(F_x(w)) - F_x(w)}{1 - F_x(w)} \right)} \\ &= \frac{1 + \xi(1 + \Gamma(x)) \left(\frac{Y_\varphi(x)}{Y_\varphi(x) + xY_\theta(x)} \right)}{1 + \left(1 + \frac{1}{\varepsilon} \right) \kappa} \\ &= \frac{1 + \xi\Gamma(x)}{1 + \left(1 + \frac{1}{\varepsilon} \right) \kappa}. \end{aligned}$$

The first equality is from (11). The second uses (iv) and (i), which implies $\eta(w) = \Psi(F_x(w)) - F_x(w)$. The third line uses (ii) to simplify the numerator, and (iii) and (v) to take the limits of the two terms in the denominator. The final step is simple algebra. The top tax rate result for the Pareto optimum follows with a little re-arranging. Setting $\xi = 0$ yields the result for the SCPE. Corollary 4 and (iv) imply $\xi > 0$, since $\Psi''(x) < 0$ implies that $g_x(w)/f_x(w) = \Psi'(F_x(w))$ is decreasing in w .