

A Theory of Income Taxation under Multidimensional Skill Heterogeneity*

Casey Rothschild
Wellesley

Florian Scheuer
Stanford

January 2014

Abstract

We develop a unifying framework for optimal income taxation in multi-sector economies with general patterns of externalities. Agents in this model are characterized by an N -dimensional skill vector corresponding to intrinsic abilities in N potentially externality-causing activities. The private return to each activity depends on individual skill and an aggregate activity-specific return, which is a fully general function of the economy-wide distribution of activity-specific efforts. We show that the N -dimensional heterogeneity can be collapsed to a one-dimensional, endogenous statistic sufficient for screening. The optimal tax schedule features a multiplicative income-specific correction to an otherwise standard tax formula. Because externalities change the relative returns to different activities, corrective taxes induce changes in the across-activity allocation of effort. These relative return effects cause the optimal correction to diverge, in general, from the Pigouvian tax that would align private and social returns. We characterize this divergence and its implications for the shape of the tax schedule both generally and in a number of applications, including externality-free economies, increasing and decreasing returns to scale, zero-sum activities such as bargaining or rent extraction, and positive or negative spillovers.

*Email addresses: crotsch@wellesley.edu, scheuer@stanford.edu. We thank Lizo Chen, Gordon Dahl, Alex Diesl, Roger Gordon, Guy Laroque, Guido Menzio, Torsten Persson, Edouard Schaal, Johannes Spinnewijn, Venky Venkateswaran, Gianluca Violante and seminar participants at Bocconi, IIES Stockholm, LSE/IFS, NYU, University of Pennsylvania, Toulouse, UC San Diego and Zurich for helpful comments. All errors are our own.

1 Introduction

How to design redistributive income tax systems is both a classic question in economics and a recurrent topic in public policy debates, as exemplified by the recent “Occupy” and “Tea Party” movements. While the standard equity-efficiency tradeoff, i.e., the tension between redistributive goals and tax distortions, which has long been emphasized by the formal optimal taxation literature,¹ has played some role, the recent debate has pointed to two central issues that have not been captured by this canonical framework. First, the trend towards greater income inequality in the past decades (as documented e.g. by Atkinson et al., 2011) has gone hand in hand with shifts in the sectoral structure of the economy, for instance a flow towards finance at the top of the income distribution. Second, supporters of the recent calls for higher taxes on high earners have questioned whether wages in some occupations actually fully reflect the true social marginal product of these activities.

Motivated by these observations, this paper provides a general framework for the analysis of optimal income taxation in multi-sector economies with endogenous wages and arbitrary patterns of externalities. In particular, individuals can pursue N different activities, the returns to each of which may depend on the aggregate efforts in this and all other activities, and in a way not necessarily aligned with marginal products. We allow for an extremely rich structure of heterogeneity, where individuals can differ along N continuous dimensions of private information, namely a skill type for each of the N activities. Tax policies in this setting reflect two key novel effects: First, sectoral shifts of effort in response to changes in the relative returns to different activities induced by changes in the income tax; and second, Pigouvian motives for taxation, correcting the wedge between wages and social returns to effort in different sectors and hence different parts of the income distribution.

Our unifying theory encompasses many applications as special cases, some of which have appeared earlier in our work. In Rothschild and Scheuer (2013b), we have considered the simplest framework for illustrating the first of the two effects above: A two-sector economy with a constant returns to scale aggregate production function and private returns equal to marginal products. With complementary sectors, the income tax schedule can be used to manipulate the relative returns to the two sectors and thereby achieve redistribution indirectly through general equilibrium effects. In Rothschild and Scheuer (2013a), we have added the second effect, again in the most parsimonious way: One of the two activities is rent-seeking and imposes negative externalities, so its private returns

¹See Mirrlees (1971), Diamond (1998), Saez (2001), Werning (2007).

exceed its social marginal product, and the second, traditional activity generates no externalities. We use this to demonstrate how the optimal correction can deviate from the partial equilibrium, Pigouvian correction due to the general equilibrium effects from sectoral shifts of effort between productive and unproductive work.

Although instructive, these examples remain restrictive in capturing many real-world settings. For instance, imagine a team production setting where individuals spend effort both to actually produce output and to claim credit (and get paid) for the output they or others have produced. Since claiming credit is a zero sum activity from a social perspective, its private returns will typically exceed its social returns. On the other hand, the some of the returns to the productive activity are captured by effort in the other activity. Hence, this is a setting where both activities generate externalities—one negative, and the other positive.

Some recent contributions to the taxation literature have addressed related phenomena. For instance, Piketty et al. (2013) have emphasized that some top incomes may come at the expense of lower incomes, e.g. because executive officers may set their compensation through bargaining, so when they claim a larger share of the resources in the company, they leave less for workers. Besley and Ghatak (2013) argue that some sectors may capture resources from other sectors, e.g. in the form of bailouts from productive workers to the financial sector. Lockwood et al. (2013) consider a model with multiple occupations, some over- and some underpaid, with different relative representations in different parts of the income distribution, justifying a purely Pigouvian role for the income tax. However, all these papers assume a very simple pattern of externalities, in the sense that whenever some activity is overpaid, this comes at the expense of everyone else *uniformly*, rather than at the expense of some more than others.

In contrast, the unifying framework we develop here allows us to consider activities that can be linked through arbitrarily rich externality structures: some activities may generate positive and others negative externalities, and the externalities may be borne *differently* by different activities. For instance, an increase in aggregate effort in the claiming credit activity in the above example clearly reduces the returns to the productive activity. But it may also reduce the return to claiming credit itself, e.g. when this activity is subject to crowding. Depending on which effect is stronger, the *relative* return to the unproductive activity may rise or fall. This in turn determines whether a marginal tax increase at incomes where the unproductive activity is strongly represented leads to a beneficial flow of effort to the productive activity, or a perverse shift to the *unproductive* activity.

These sectoral shift effects in response to relative return changes turn out to play an important role for optimal tax policy. We derive a useful formula for our general frame-

work that offers insight into the size and direction of the divergence between the optimal correction and the partial equilibrium Pigouvian correction that ignores these relative return effects. We also show that this divergence vanishes precisely when a variation in the marginal income tax rate at a given income level induces no relative return changes. We use these general results in various specific applications to characterize both the optimal level and the optimal progressivity of the income tax schedule for any redistributive objectives, captured by arbitrary Pareto weights.

Since our model naturally involves N dimensions of private information, we begin by demonstrating how they can be collapsed into a single dimension relevant for screening, extending our previous work in Rothschild and Scheuer (2013a,b).² Although settings with multidimensional heterogeneity are frequently challenging to solve (Rochet and Choné, 1998), we identify a one-dimensional, but endogenous, summary statistic for heterogeneity in our framework. Because taxes can only condition on income, not on how it is earned through different activities, an individual always earns a given amount of income through a cost-minimizing combination of efforts in the N activities. For any vector of activity-specific returns, this results in a well-defined *wage* that determines her preferences over consumption-income bundles. We can therefore work with a screening problem in terms of these wages, with the only complication that they depend on sectoral returns and therefore the vector of aggregate efforts in all activities.

We first solve this screening problem for any given combination of sectoral efforts (the “inner” problem), which allows us to obtain a formula for the marginal income tax rate in any Pareto optimum (Proposition 1). It closely mirrors the formula for a standard Mirrlees model, but features an additional adjustment factor that captures the optimal correction both for externalities and relative return effects. The remainder of the paper is then focused on characterizing precisely this adjustment factor. Since this is closely related to finding the optimal combination of aggregate efforts in each activity for a given set of Pareto weights (the “outer” problem), we describe the welfare effects of marginal variations in these efforts in some detail, which prominently feature the sectoral shift effects that we emphasize (Lemma 4).

We then use the resulting optimality condition to characterize the adjustment factor in the marginal tax rate formula and, more importantly, compare it to the partial equilibrium, Pigouvian correction, which is simply the income share weighted average, at each income level, of the wedges between the private returns and social marginal products of the activities, as e.g. in Lockwood et al. (2013). Proposition 2 shows that the two coincide

²For other recent studies of optimal taxation under multidimensional heterogeneity, see e.g. Kleven et al. (2009), Choné and Laroque (2010) and Scheuer (2013).

precisely at income levels where a variation in the marginal tax rate has no relative return effects. Based on this, Proposition 3 provides conditions under which the dimensionality of the Pareto problem can be reduced: If there are K directions in the space \mathbb{R}^N of aggregate sectoral effort vectors in which there are both no relative return effects and no externalities, then the outer problem effectively collapses to an $N - K$ -dimensional problem with $N - K$ consistency constraints. We identify both special cases in Rothschild and Scheuer (2013a,b) as applications of this general principle, where two-sector models can be solved with a single sufficient statistic for the wage distribution.

We then illustrate how the general tools developed here can be used to characterize the shape of optimal tax schedules in a number of important applications, two of which extend our earlier work and the rest of which are novel. The first investigates how the results from the externality-free environment with two sectors considered in Rothschild and Scheuer (2013b) extend to settings with more than two sectors (Proposition 4). We argue that the presence of additional sectors can reinforce the regressive adjustment to the standard Mirrleesian tax schedule that we found for the two-sector model, effectively moving the optimal progressivity of the income tax closer to that in a model with fixed occupations, such as Stiglitz (1982).

The second application adds aggregate externalities in the form of increasing or decreasing returns to scale to the two-sector model. In this case, the adjustment factor can be transparently decomposed into a local and global component (Proposition 5). The first, which depends on the income shares of the two activities at any given income level, has the same regressive form as in the no externalities case, capturing relative return effects. The second, uniform across income levels, accounts for the externalities and simply scales all marginal tax rates up (down) under decreasing (increasing) returns to scale.

We then consider the case where aggregate technology exhibits constant returns to scale, but the sectoral income shares are decoupled from marginal products, as motivated by the example discussed at the beginning (Proposition 6). For instance, suppose the relatively high-wage activity is also the overpaid activity, in the sense that its aggregate income share exceeds what would correspond to its marginal product. Then the Pigouvian correction implies a *more* progressive income tax schedule than in a standard Mirrlees model. However, the optimal correction exceeds this Pigouvian correction at any given income level if an increase in the marginal tax rate reduces the relative return to the overpaid activity and thus induces a beneficial shift of effort out of it, and vice versa.

Finally, we turn to two applications that we can fully characterize for general N , namely the case where all returns depend only on the aggregate effort in one activity (Proposition 7), or where the returns to all activities are fixed, except for one, which

depends on the aggregate efforts in all activities (Proposition 8). The first is a generalization of Rothschild and Scheuer (2013b), allowing for more than one traditional activity and positive or mixed externalities, such as positive spillovers from research or entrepreneurial activities onto other sectors, but within-sector crowding effects.

By restricting attention to taxes on income, rather than instruments that perfectly target its decomposition into earnings from specific activities, our paper crucially differs from most of the literature on corrective taxation in the presence of externalities.³ This restriction prevents the simple “principle of targeting” (Dixit, 1985) from applying, under which the optimal correction is equal to the Pigouvian tax, and therefore makes our analysis both theoretically interesting and practically relevant. Instead, we show how, under imperfect instruments, the adjustment to the optimal tax formula diverges from the Pigouvian tax depending on relative return effects.

Our analysis is more closely related to Diamond (1973), who shows that the optimal linear tax on an externality producing consumption good can be expressed as a correction capturing the direct effect on the demand for the good, and an adjustment term that reflects the indirect, general equilibrium effect of the changes in consumption of the good induced by the direct effect. We have a distinct motivation for such general equilibrium effects, namely in the form of effort choice along various activities, which allows us to provide insights into how the optimal correction deviates from the Pigouvian tax. Moreover, whereas Diamond’s analysis is in the Ramsey framework, we focus on nonlinear taxation in a Mirrlees setting, including both redistributive and corrective motives for taxation.⁴

The paper is organized as follows. Section 2 introduces the model, provides some simple illustrations of its flexibility, and shows how the multidimensional screening problem can be collapsed. Section 3 provides the general N -sector results, including the marginal tax rate formula and the key optimality conditions for the outer problem. Section 4 provides a further characterization for $N = 2$ and Section 5 collects the discussion of the applications. All proofs are relegated to an appendix.

2 The Model

2.1 Setup

We consider an economy in which individuals can pursue N different activities, indexed by i . Each agent is characterized by an N -dimensional unobservable skill vector $\theta \in \Theta \equiv$

³See Sandmo (1975), Sadka (1978), Cremer et al. (1998) and Kopczuk (2003).

⁴Similarly to Diamond (1973), Lockwood et al. (2013) also abstract from redistributive concerns.

$\prod_{i=1}^N \Theta_i$, where the i^{th} element $\theta_i \in \Theta_i = [\underline{\theta}_i, \bar{\theta}_i]$ captures her skill in activity i . We assume $\theta_i > 0$ for all i . Skills are distributed with a continuous N -dimensional cdf $F : \Theta \rightarrow [0, 1]$ and associated pdf $f(\theta)$.

Individual preferences are characterized by a continuously differentiable and concave utility function over consumption c and the vector of efforts in each activity, $e = (e_1, \dots, e_N)$, given by $U(c, e) = u(c, m(e)) \equiv u(c, l)$. We assume $u_c > 0$, $u_l < 0$, and that the effort aggregator $m(e)$ is increasing in all arguments, continuously differentiable, strictly quasiconvex and linear homogeneous.⁵ We denote the consumption and vector of activity-specific efforts of an individual of type θ by $c(\theta)$ and $e(\theta) = (e_1(\theta), \dots, e_N(\theta))$, and the total individual effort and utility by $l(\theta) \equiv m(e(\theta))$ and $V(\theta) \equiv u(c(\theta), l(\theta))$.

Aggregate output (and hence income) $Y(E)$ consists of the aggregate incomes generated in each activity $Y^i(E)$, so $Y(E) = \sum_{i=1}^N Y^i(E)$, where

$$E_i \equiv \int_{\Theta} \theta_i e_i(\theta) dF(\theta) \quad (1)$$

is the aggregate effective (i.e., skill-weighted) effort in activity i , and each Y^i can depend on the entire vector of aggregate efforts $E \equiv (E_1, \dots, E_N)$. The income of an individual of type θ in activity i is $y_i(\theta)$, and her total income from all activities is $y(\theta) \equiv \sum_{i=1}^N y_i(\theta)$. Accordingly, aggregate total and sectoral incomes are $Y(E) = \int_{\Theta} y(\theta) dF(\theta)$ and $Y^i(E) = \int_{\Theta} y_i(\theta) dF(\theta)$ for all i .

At this point, we remain fully general about the form of technology linking the N sectors, with the only assumption that each unit of effective effort in a given sector has the same private return. Formally, for each activity i , there exists some return $r_i(E)$ such that $y_i(\theta) = r_i(E) \theta_i e_i(\theta)$ for all $\theta \in \Theta$. As a result, using (1), $Y^i(E) = r_i(E) E_i$ and we can write $Y(E) = \sum_{i=1}^N r_i(E) E_i$. Note that the returns r_i may deviate from the marginal product of effort in activity i , given by $Y_i(E) \equiv \partial Y(E) / \partial E_i$, thus allowing for arbitrary patterns of externalities as illustrated next.

2.2 Examples

A simple example occurs when $Y(E)$ is a standard neoclassical production function with $r_i(E) = Y_i(E)$ for all E , so returns correspond to marginal products. For instance, in the

⁵This allows for preferences $\tilde{u}(c, \tilde{m}(e))$ where \tilde{m} is homothetic even if not linear homogeneous: then there exists some increasing $h(\cdot)$ and linear homogeneous $m(e)$ such that we can write $\tilde{m}(e) = h(m(e))$, and we can define $u(c, l) \equiv \tilde{u}(c, h(l))$. Hence, $\tilde{u}(c, \tilde{m}(e)) = u(c, m(e))$ for all (c, e) . An example is $\hat{u}(c) - \sum_{i=1}^N h_i(e_i)$ when all $h_i(\cdot)$ are homogeneous of the same degree. A limiting case would obtain for $m(e) = \sum_{i=1}^N e_i$, so that m is weakly but not strictly quasiconvex. Then individuals would always specialize in the activity that yields them the highest returns, as in the Roy models considered in Rothschild and Scheuer (2011, 2013).

limiting case where $m(e)$ becomes linear and individuals always specialize in one of the N activities, we could interpret $Y(E)$ as an aggregate production function for an economy with N complementary sectors or occupations i , and characterize an optimal income tax in such a multi-sector Roy (1951) model, as in Rothschild and Scheuer (2013b) for $N = 2$.

The recent debate about tax policy has started to question whether private and social returns always coincide. For instance, incomes in some activities may come at the expense of others, through bargaining, rent-seeking or negative externalities. Consider e.g. the team production setting from the introduction where individuals spend time and effort both to produce output (activity 2) and to get credit (and compensated) for this output (activity 1). This can be captured by $Y(E) = E_2$ and $Y^1(E) = a(E_1)E_2$, $Y^2(E) = (1 - a(E_1))E_2$, where $a(E_1)$ is some increasing function. Here, activity 2 generates positive externalities as it increases the returns $r_1 = a(E_1)E_2/E_1$ to activity 1, and activity 1 imposes negative externalities on activity 2. For instance, in Biais et al. (2011), fast traders impose externalities on slow traders through adverse selection from their information advantage. Glode and Lowery (2012) consider a model where financial sector workers engage in both (unproductive) speculative trading and surplus creation (e.g. from market making) and argue that profits from both activities are interlinked.

Another example for a pure zero sum activity would be a setting where activity 1 just takes away output produced through activity 2 one-for-one, so that $Y(E) = Y(E_2)$ and $Y^1(E_1) = E_1$, $Y^2(E) = Y(E_2) - E_1$. Here, both activities again generate externalities, but only on the returns $r_2(E) = (Y(E_2) - E_1)/E_2$ to the productive activity 2 (the returns to activity 1 are fixed at 1, so it bears no externalities). The opposite special case is considered in Rothschild and Scheuer (2013a) (again for $N = 2$), where only one (rent-seeking) activity imposes (negative) externalities on itself and all other activities, so $r_i(E_1)$ for all i and all r_i are decreasing. This could capture negative externalities from search activities with crowding effects, e.g. for profitable arbitrage opportunities in financial markets, or tournaments and races with winner-takes-all compensation in the arts, entertainment, law or R&D. On the other hand, our general framework can also allow for positive externalities, such as spillover effects from entrepreneurial and innovative activities.

These examples make clear that individual activities may be hard to target separately. As emphasized in Lockwood et al. (2013), it may also be the case that some identifiable sectors or professions are on average over- or underpaid. Our analysis formally applies in these settings as well. Of course, separate, occupation-specific tax schedules—which we do not consider here—will typically be desirable in these settings. Our analytical framework is still useful insofar as implementability problems, including relabeling of activities when there is a “fuzzy frontier,” special interest lobbying for lower taxes, horizontal eq-

uity concerns, or multiple activities within professions (in which case our analysis applies even to a profession-specific tax), render such targeted taxes impractical. We revisit the above and other examples in Section 5.

2.3 Income Tax Implementation

We first describe the set of feasible allocations using a direct mechanism and then link this to the implementation through a nonlinear income tax schedule. Individuals announce their type θ and then get assigned consumption $c(\theta)$, total income $y(\theta)$, and the fraction of income earned through each activity i , given by $q^i(\theta) \equiv y_i(\theta)/y(\theta) = r_i(E)\theta_i e_i(\theta)/y(\theta)$. Let $q(\theta) \equiv (q^1(\theta), \dots, q^N(\theta)) \in \Delta^{N-1}$ be the vector of these income shares, where $\Delta^{N-1} = \{q \in \mathbb{R}^N \mid \sum_{i=1}^N q_i = 1, q_i \geq 0\}$ is the $N - 1$ dimensional unit simplex.

Assuming that only income y and consumption c are observable, but not an individual's skill type θ nor their sectoral efforts e_i (nor the income shares q^i), the incentive constraints that guarantee truth-telling of the agents are:

$$\begin{aligned} & u \left(c(\theta), m \left(\frac{q^1(\theta)y(\theta)}{\theta_1 r_1(E)}, \dots, \frac{q^N(\theta)y(\theta)}{\theta_N r_N(E)} \right) \right) \\ & \geq \max_{p \in \Delta^{N-1}} \left\{ u \left(c(\theta'), m \left(\frac{p_1 y(\theta')}{\theta_1 r_1(E)}, \dots, \frac{p_N y(\theta')}{\theta_N r_N(E)} \right) \right) \right\} \quad \forall \theta, \theta' \in \Theta, \end{aligned} \quad (2)$$

since each type θ can imitate type θ' by earning the income of type θ' using a continuum of efforts combinations and hence income shares $p = (p_1, \dots, p_N)$ in the N activities.

The following two results show how the N -dimensional incentive constraints (2) can be collapsed into a one-dimensional screening problem in terms of an endogenous summary statistic of heterogeneity. First, incentive compatibility implies that each type θ has a well-defined wage $w \equiv y/l$ and vector of sectoral income shares q , which both depend on E but are otherwise independent of the allocation.

Lemma 1. *Suppose that only income y and consumption c are observable, whereas an individual's skill type θ and vector of efforts e and income shares q are private information. Then, in any incentive compatible allocation $\{c(\theta), y(\theta), q(\theta), E\}$,*

$$w(\theta) \equiv \frac{y(\theta)}{l(\theta)} = \max_{p \in \Delta^{N-1}} m \left(\frac{p_1}{\theta_1 r_1(E)}, \dots, \frac{p_N}{\theta_N r_N(E)} \right)^{-1} \quad (3)$$

and $q(\theta)$ is the corresponding arg max.

Lemma 1, which generalizes the result for $N = 2$ in Rothschild and Scheuer (2013a), establishes that, in any incentive compatible allocation, each type's wage $w(\theta)$ is fully

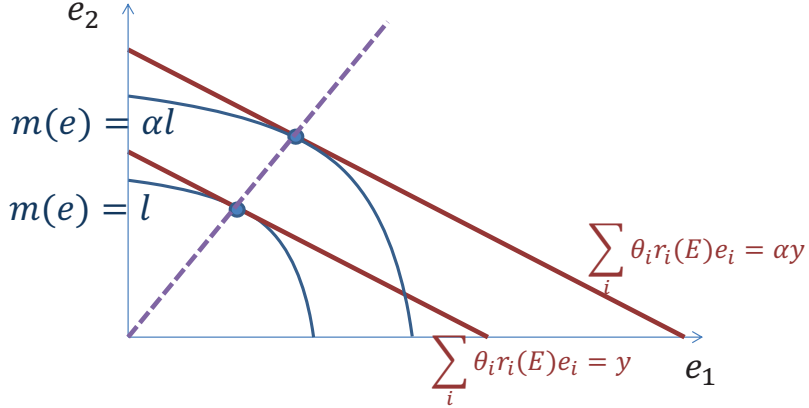


Figure 1: Determination of wages and income shares given E

pinned down by the vector E . To make this explicit, we write $w_E(\theta)$ in the following. Moreover, the vector of income shares $q(\theta)$ is chosen so as to minimize the overall effort $m(e)$ subject to achieving a given amount of income: By (3) and linear homogeneity of m ,

$$w_E(\theta) = \max_{p \in \Delta^{N-1}} y m \left(\frac{p_1 y}{\theta_1 r_1(E)}, \dots, \frac{p_N y}{\theta_N r_N(E)} \right)^{-1} = \max_e \frac{y}{m(e)} \text{ s.t. } \sum_{i=1}^N \theta_i r_i(E) e_i = y \quad (4)$$

for any y . Again by linear homogeneity of m , the vector $q(\theta)$ only depends on the vector of skill ratios $\phi \equiv (\theta_1/\theta_N, \dots, \theta_{N-1}/\theta_N)$ in addition to E , which is why we write $q_E(\phi)$ henceforth, with $\phi \in \Phi \equiv (0, \infty)^{N-1}$. Since m is strictly quasiconvex, $q_E(\phi)$ is uniquely determined by E for each ϕ .

Figure 1 illustrates the intuition underlying Lemma 1 for the case of two activities. By (4), individuals choose their efforts e_1 and e_2 to minimize their overall effort $m(e)$ subject to achieving a given amount of income y . When the targeted amount of income changes, say by factor α , their optimal effort ratio e_1/e_2 , and $m(e)$, also increase by the factor α due to the linear homogeneity of m . Hence, their income shares q^1 and q^2 as well as their wage $y/m(e)$ are independent of y and only depend on the slope of the lines in Figure 1, which is fully pinned down by the returns $r_i(E)$.

All individuals with the same wage w have the same preferences over (c, y) -bundles given by $u(c, y/w)$. As is standard, we assume the single crossing property, i.e., that the marginal rate of substitution between y and c , $-u_l(c, y/w) / (w u_c(c, y/w))$, is decreasing in w . Under this condition, any incentive compatible allocation can be implemented with a non-linear income tax $T(y)$.

Lemma 2. Consider the observability assumptions from Lemma 1 and suppose that the allocation

$\{c(\theta), y(\theta), q(\theta), E\}$ is incentive compatible. Then

(i) $w_E(\theta) = w_E(\theta') = w$ implies $u(c(\theta), y(\theta)/w) = u(c(\theta'), y(\theta')/w)$, and

(ii) $\{c(\theta), y(\theta), q(\theta), E\}$ can be implemented by offering the single non-linear income tax schedule $T^*(y)$ corresponding to the retention function $R^*(y) = y - T^*(y)$ defined by

$$R^*(y) \equiv \max_c \left\{ c \mid u \left(c(\theta), \frac{y(\theta)}{w_E(\theta)} \right) \geq u \left(c, \frac{y}{w_E(\theta)} \right) \forall \theta \in \Theta \right\} \quad (5)$$

and letting all agents choose one of their most preferred (c, y) -bundles from the resulting budget set $B^* = \{(c, y) \mid c \leq y - T^*(y)\}$.

Lemma 2 does not rule out the possibility that two individuals with the same w (but different q 's) choose different (c, y) -bundles, even though, by property (i), these bundles must be located on the same indifference curve in the (c, y) -space. To trace out the Pareto frontier, we can nevertheless restrict attention to allocations $\{c(w), y(w), E\}$ that pool all same-wage individuals at the same (c, y) -bundle by the arguments in Rothschild and Scheuer (2013b). We focus on such allocations in the following.

3 N Sectors

3.1 Definitions

We use general cumulative Pareto weights $\Psi(\theta)$ defined over the N -dimensional Θ -space with the corresponding density $\psi(\theta)$ to trace out the set of Pareto efficient allocations. The social planner maximizes $\int_{\Theta} V(\theta) d\Psi(\theta)$ subject to resource and self-selection constraints. The observation that makes this problem tractable is that, by Lemma 1, fixing the vector E determines the wage $w_E(\theta)$ and the vector of income shares $q_E(\phi)$ for each type θ . For any given E , the cdf over (w, ϕ) -vectors, given by

$$G_E(w, \phi) \equiv \int_{\{\theta \mid w_E(\theta) \leq w, \theta_i/\theta_N \leq \phi_i \forall i=1, \dots, N-1\}} dF(\theta)$$

with the corresponding density $g_E(w, \phi)$ will be useful in the following. It is useful precisely because all individuals who earn the same wage w , and are therefore located on the same iso-wage curve in θ -space as drawn in Figure 2, are pooled in the same allocation. Nonetheless, these individuals differ in their relative skills ϕ through which they achieve wage w . Hence, conditional on E , each θ -type is fully identified by her (w, ϕ) -combination. We denote the support of the wage distribution for any E by $[\underline{w}_E, \bar{w}_E]$, where $\underline{w}_E = w_E(\underline{\theta}_1, \dots, \underline{\theta}_N)$ and $\bar{w}_E = w_E(\bar{\theta}_1, \dots, \bar{\theta}_N)$. This allows us to obtain the wage

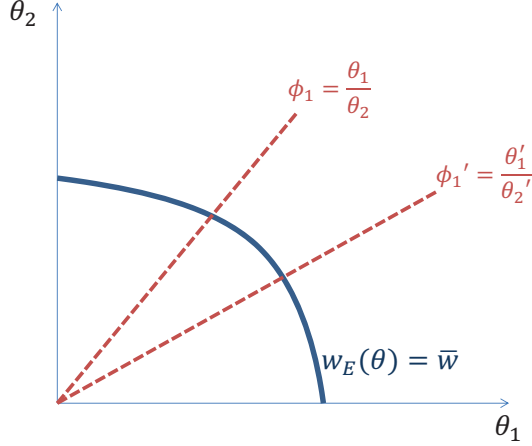


Figure 2: Pooling along iso-wage curves in θ -space conditional on E

distribution for any given E simply as the marginal distribution

$$F_E(w) \equiv \int_{\{\theta | w_E(\theta) \leq w\}} dF(\theta) = \int_{\underline{w}_E}^w \int_{\Phi} dG_E(z, \phi)$$

with the corresponding density $f_E(w) = \int_{\Phi} dG_E(w, \phi)$ as well as the sectoral densities $f_E^i(w) \equiv \int_{\Phi} q_E^i(\phi) dG_E(w, \phi)$. Hence, the sectoral densities can be interpreted as an average value of q^i for all individuals at wage w , and $f_E(w) = \sum_{i=1}^N f_E^i(w)$ for all $w \in [\underline{w}_E, \bar{w}_E]$.⁶ Finally, for any given E , we can derive Pareto weights over wages $\Psi_E(w)$, as well as their density and sectoral decomposition $\psi_E(w) = \sum_{i=1}^N \psi_E^i(w)$, completely analogously from $\Psi(\theta)$.

By the discussion following Lemma 2, we can focus on incentive compatible allocations $\{c(w), y(w), E\}$ that only condition on an individual's wage w , which then imply total effort and utility $l(w) \equiv y(w)/w$ and $V(w) \equiv u(c(w), l(w))$ as well as the activity-specific efforts $e_i(\theta) = q_E^i(\phi)y(w_E(\theta))/(\theta_i r_i(E))$.

3.2 Inner and Outer Problems for Pareto Efficiency

As in Rothschild and Scheuer (2013a,b), we decompose the problem of finding Pareto optimal allocations into two steps. The first step involves finding the optimal vector of aggregate efforts E . We call this the “outer” problem. The second (which we call the “inner” problem) involves finding the optimal resource-feasible and incentive-compatible allocation for a given E . This inner problem is an almost standard Mirrlees problem; the

⁶In the limiting case with $m(e) = \sum_{i=1}^N e_i$, (3) immediately implies $q_E^i(\phi) \in \{0, 1\}$ and $w_E(\theta) = \max\{\theta_1 r_1(E), \dots, \theta_N r_N(E)\}$. Then $f_E^i(w)/f_E(w)$ can be interpreted as the share of i -sector workers at w , whereas here it is the i -sector income share at wage w .

only difference is that the induced vector of aggregate effective efforts has to be consistent with the E that we are fixing for the inner problem. For some given Pareto weights $\Psi(\theta)$ (and hence induced weights $\Psi_E(w)$), we therefore define the inner problem as follows (where $c(V, l)$ is the inverse function of $u(c, l)$ w.r.t. c):

$$W(E) \equiv \max_{V(w), l(w)} \int_{\underline{w}_E}^{\bar{w}_E} V(w) d\Psi_E(w) \quad (6)$$

subject to
$$V'(w) = u_l(c(V(w), l(w)), l(w)) \frac{l(w)}{w} \quad \forall w \in [\underline{w}_E, \bar{w}_E] \quad (7)$$

$$E_i = \frac{1}{r_i(E)} \int_{\underline{w}_E}^{\bar{w}_E} w l(w) f_E^i(w) dw \quad \forall i = 1, \dots, N \quad (8)$$

$$\int_{\underline{w}_E}^{\bar{w}_E} w l(w) f_E(w) dw \geq \int_{\underline{w}_E}^{\bar{w}_E} c(V(w), l(w)) f_E(w) dw. \quad (9)$$

We employ the standard Mirrleesian approach of optimizing directly over allocations, i.e., over effort $e(w)$ and consumption or, equivalently, utility $V(w)$ profiles. The social planner maximizes a weighted average of individual utilities $V(w)$ subject to three sets of constraints. (9) is a standard resource constraint. The N constraints in (8) ensure that aggregate effective effort in each sector i indeed sums up to E_i , as the right-hand-side is

$$\frac{1}{r_i(E)} \int_{\underline{w}_E}^{\bar{w}_E} y(w) f_E^i(w) dw = \int_{\underline{w}_E}^{\bar{w}_E} \int_{\Phi} \frac{y(w) q_E^i(\phi)}{r_i(E)} g_E(w, \phi) d\phi dw = \int_{\Theta} \theta_i e_i(\theta) dF(\theta). \quad (10)$$

Finally, the allocation $V(w), l(w)$ needs to be incentive compatible, i.e.,

$$V(w) \equiv u(c(w), l(w)) = \max_{w'} u \left(c(w'), \frac{l(w')w'}{w} \right). \quad (11)$$

It is a well-known result that under single-crossing, the global incentive constraints (11) are equivalent to the local incentive constraints (7) and the monotonicity constraint that income $y(w)$ must be non-decreasing in w .⁷ We follow the standard approach of dropping the monotonicity constraint, which could easily be checked ex post. If the solution to problem (6) to (9) does not satisfy it, bunching would need to be considered.

Once a solution $V(w), l(w)$ to the inner problem has been found, the resulting welfare is given by $W(E)$. The outer problem is then simply $\max_E W(E)$.

Solving the inner problem (6) to (9) for a given E yields the following optimal marginal tax rate formula:

⁷See, for instance, Fudenberg and Tirole (1991), Theorems 7.2 and 7.3.

Proposition 1. *The marginal tax rate in any Pareto optimum without bunching is such that*

$$1 - T'(y(w)) = \left(1 - \sum_{i=1}^N \frac{\xi_i}{r_i(E)} \frac{f_E^i(w)}{f_E(w)}\right) \left(1 + \frac{\eta(w)}{w f_E(w)} \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)}\right)^{-1} \quad \text{with} \quad (12)$$

$$\eta(w) = \int_w^{\bar{w}_E} \left(1 - \frac{\Psi_E(s)}{f_E(s)} \frac{u_c(s)}{\lambda}\right) \exp\left(\int_w^s \left(1 - \frac{\varepsilon^u(t)}{\varepsilon^c(t)}\right) \frac{dy(t)}{y(t)}\right) f_E(s) ds \quad (13)$$

for all $w \in [\underline{w}_E, \bar{w}_E]$, where λ is the multiplier on the resource constraint (9), $\lambda \xi_i$ are the multipliers on the N consistency constraints (8), $\lambda \hat{\eta}(w) = \lambda \eta(w) / u_c(w)$ the multipliers on the local incentive constraints (7), and $\varepsilon^c(w)$ ($\varepsilon^u(w)$) the (un)compensated wage elasticity of total effort l .

These formulas closely mirror the formulas in a standard Mirrlees model (see e.g. equations (15) to (17) in Saez, 2001). The term $\eta(w)$ captures the redistributive motives of the government and income effects from the terms in the exponential function. This simplifies with quasilinear preferences $u(c, l) = c - h(l)$, where income effects disappear, as in Diamond (1998). Then $u_c(w) = \lambda = 1$ and $\varepsilon^u(w) = \varepsilon^c(w) \forall w$, so that $\eta(w) = \Psi_E(w) - F_E(w)$. Hence the marginal tax rate is increasing in the degree to which $\Psi_E(w)$ shifts weight to low-wage individuals compared to $F_E(w)$.

The only difference from standard formulas is that, at each wage, the marginal keep shares $1 - T'(y(w))$ are adjusted by a correction factor $1 - \sum_{i=1}^N (f_E^i(w) / f_E(w)) (\xi_i / r_i(E))$. As we discuss further below, this factor is a local correction for the general equilibrium effects and/or externalities caused by income earned by wage w individuals. In particular, the multiplier ξ_i on the i^{th} constraint (8) is the *general equilibrium* Pigouvian corrective tax on effective effort in sector i —i.e., the optimal corrective tax taking general equilibrium effects into account. The term $\sum_{i=1}^N (f_E^i(w) / f_E(w)) (\xi_i / r_i(E))$ is therefore an income-share weighted average of the general equilibrium Pigouvian corrective taxes ξ_i / r_i on the *incomes* earned in the various activities.

We next use the conditions for an optimal E from outer problem to explore the relationship between the general equilibrium corrections ξ_i / r_i and the *partial equilibrium* Pigouvian taxes τ_p^i that would align the social and private marginal products of effort in sector i , defined by $r_i(E)(1 - \tau_p^i(E)) \equiv Y_i(E)$.

3.3 Outer Problem

At the optimal E from the outer problem, there must be zero net welfare effects from a marginal change in any E_i . We can divide the marginal welfare effects of such a change into four classes: (i) the direct effect on the left-hand side of the i^{th} consistency constraint (8) and three other effects, (ii)-(iv), which are best understood by considering the effect

of a marginal change in E_i on any given type θ . The change in E_i changes the wage of individual i . We designate by (ii) the direct effects that this wage change has on (6) to (9), holding fixed the type's effort and utility. We designate by (iii) the indirect effects that this wage change has on θ 's effort-utility allocation as she moves along the fixed schedules $l(w), V(w)$ when her wage changes. Finally, the change in E_i also changes the returns $r_i(E)$ to effort in the various sectors; in turn, this changes θ 's optimal allocation of efforts $e_i(\theta)$ across the various sectors for any given total effort $l(w)$. We designate by (iv) the welfare effects of this effort-composition change on the right-hand side of the consistency constraints (8).

One approach would be to compute these effects (in terms of the multipliers on the constraints) using the envelope theorem and holding the schedules $l(w), V(w)$ fixed. A more useful alternative, pursued in the following, is to simultaneously vary the schedules $l(w), V(w)$ in way that undoes the change in average effort and utility at each w coming from (iii). In particular, note that (4) can equivalently be written as

$$w_E(\theta) = \max_e \frac{\sum_{i=1}^N \theta_i r_i(E) e_i}{m(e)} \text{ s.t. } m(e) = l. \quad (14)$$

Using the envelope theorem and denoting the semi-elasticities of the returns $r_j(E)$ in activity j w.r.t. E_i by

$$\beta_i^j(E) \equiv \frac{\partial r_j(E)}{\partial E_i} \frac{1}{r_j(E)},$$

the semi-elasticity of wages w.r.t. E_i is

$$\frac{\partial w_E(\theta)}{\partial E_i} \frac{1}{w_E(\theta)} = \frac{\sum_{j=1}^N \theta_j e_j(\theta) r_j(E) \beta_i^j(E)}{w_E(\theta) l} = \sum_{j=1}^N q_E^j(\phi) \beta_i^j(E), \quad (15)$$

i.e., the income-share weighted average of the return semi-elasticities. The change in effort for individuals with original wage w and income share vector q due to the wage change resulting from a marginal increase dE_i in E_i is therefore $l'(w) w \sum_{j=1}^N q_E^j \beta_i^j(E) dE_i$ and the average effort change at w is simply

$$l'(w) w \sum_{j=1}^N \mathbb{E} \left[q_E^j(\phi) \mid w \right] \beta_i^j(E) dE_i = l'(w) w \sum_{j=1}^N \frac{f_E^j(w)}{f_E(w)} \beta_i^j(E) dE_i, \quad (16)$$

where $\mathbb{E}[q_E^j(\phi) \mid w] = \int_{\Phi} q_E^j(\phi) g_E(\phi \mid w) d\phi$ is the average of q^j over the set $\{\theta \mid w_E(\theta) = w\}$

of all wage- w individuals. Defining

$$\delta_E^i(w) \equiv \sum_{j=1}^N \frac{f_E^j(w)}{f_E(w)} \beta_i^j(E), \quad (17)$$

this motivates the variation in the l -schedule $\tilde{l}(w) = l(w) - l'(w)w\delta_E^i(w)dE_i$. Analogously, we can vary the V -schedule by $\tilde{V}(w) = V(w) - V'(w)w\delta_E^i(w)dE_i$. Performing this variation in schedules simultaneously with the increase in E_i greatly simplifies the outer problem effects (iii) by making sure that both average effort and utility for the set of types at each wage w remain unchanged when E_i increases marginally. In fact, this variation also ensures that average consumption is unchanged at each w when E_i increases.⁸

We refer below to subshift 1 as the change in the total effort $l(w)$ and utility $V(w)$ of types originally at wage w that is due to the change in wages coming from the marginal increase in E_i , holding the l - and V -schedules fixed (i.e., effect (iii) from above). In the additional subshift 2, we move the schedules from $l(w), V(w)$ to $\tilde{l}(w), \tilde{V}(w)$. This additional subshift does not alter the welfare effect of a marginal change in E_i , since the variation in schedules has a zero total welfare effect by the envelope theorem if $l(w)$ and $V(w)$ are a solution to the inner problem.

3.3.1 Redistributive Effects

The only effects on the objective come from (iii). The effect from subshift 1 at each w is (by analogy to (16)) $V'(w)w \sum_{j=1}^N \psi_E^j(w) \beta_i^j(E)$. The effect of the variation in the V -schedule from subshift 2 is simply $-V'(w)w\delta_E^i(w)$. Taking these together, the total effect is

$$\sum_{j=1}^N \beta_i^j(E) \int_{\underline{w}_E}^{\bar{w}_E} V'(w)w \left(\frac{\psi_E^j(w)}{\psi_E(w)} - \frac{f_E^j(w)}{f_E(w)} \right) \psi_E(w)dw \equiv -\lambda \sum_{j=1}^N \beta_i^j(E) R_j(E) \quad (18)$$

with

$$R_j(E) \equiv \int_{\underline{w}_E}^{\bar{w}_E} \frac{V'(w)w}{\lambda} \left(\frac{f_E^j(w)}{f_E(w)} - \frac{\psi_E^j(w)}{\psi_E(w)} \right) \psi_E(w)dw. \quad (19)$$

Note that $\sum_{j=1}^N R_j(E) = 0$; this is because the effect captures a re-allocation of utility across workers with different sectoral intensities q at each w . For the same reason, each R^j disappears if we put the same welfare weight on all types θ who earn the same wage w (so that $\psi_E^j(w)/\psi_E(w) = f_E^j(w)/f_E(w)$ for all j, w , as would be the case with the relative weights $\Psi(\theta) = \tilde{\Psi}(F(\theta))$). Otherwise, if a marginal increase in E_i increases the returns to

⁸To wit, dropping the common argument w and using (7) and (50) yields $\tilde{c} - c = c(\tilde{V}, \tilde{l}) - c(V, l) = \frac{1}{u_c}(\tilde{V} - V) - \frac{u_l}{u_c}(\tilde{l} - l) = \left(\frac{1}{u_c} V' - \frac{u_l}{u_c} l' \right) w \delta_E^i = \left(-\frac{u_l}{w u_c} + \frac{u_c c' + u_l l / w}{u_c} \right) w \delta_E^i = c' w \delta_E^i$.

activities in which workers with a high relative welfare weight earn much of their income, then the resulting re-allocation in utilities is welfare enhancing.

3.3.2 Incentive Constraint Effects

Again, the only effects here are from (iii). There are no effects from subshift 1, since individuals just move along incentive compatible schedules. For subshift 2, note that

$$\tilde{V}'(w) = V'(w) - \frac{d(V'(w)w)}{dw} \delta_E^i(w) dE_i - V'(w)w \delta_E^{i'}(w) dE_i,$$

so the change in the local incentive constraints (7) is

$$\begin{aligned} & \tilde{V}'(w) - V'(w) + u_l(\tilde{c}(w), \tilde{l}(w)) \frac{\tilde{l}(w)}{w} - u_l(c(w), l(w)) \frac{l(w)}{w} \\ &= -\frac{d(V'(w)w)}{dw} \delta_E^i(w) dE_i - V'(w)w \delta_E^{i'}(w) dE_i - \frac{d(u_l(c(w), l(w))l(w))}{dw} \delta_E^i(w) dE_i \\ &= -V'(w)w \delta_E^{i'}(w) dE_i \end{aligned}$$

since (7) requires $wV'(w) + u_l(w)l(w) = 0$ for all w . Using (17), the incentive effects from (iii) are therefore

$$-\sum_{j=1}^N \beta_j^j(E) \lambda \int_{\underline{w}_E}^{\bar{w}_E} \eta(w) w \frac{V'(w)}{u_c(w)} \frac{d}{dw} \left(\frac{f_E^j(w)}{f_E(w)} \right) dw \equiv -\lambda \sum_{j=1}^N \beta_j^j(E) I_j(E), \quad (20)$$

where $\lambda \hat{\eta}(w) = \lambda \eta(w) / u_c(w)$ is the multiplier on (7) and

$$I_j(E) \equiv \int_{\underline{w}_E}^{\bar{w}_E} \eta(w) w \frac{V'(w)}{u_c(w)} \frac{d}{dw} \left(\frac{f_E^j(w)}{f_E(w)} \right) dw. \quad (21)$$

As before, we have $\sum_{j=1}^N I_j(E) = 0$. To interpret the terms I_j , suppose $\eta(w) > 0$, so the incentive constraints bind downwards. Then a marginal increase in E_i is welfare reducing (respectively, increasing) if it increases (decreases) the returns to activities j with $d(f_E^j(w)/f_E(w))/dw > 0$, i.e., to activities that are locally associated with high wages in the first place. This is because an increase in E_i makes the wage distribution more (less) unequal in this case, which tightens (loosens) the local incentive constraints. The effect is therefore a generalized version of the one pointed out by Stiglitz (1982) for a two-type model with two sectors.

Figure 3 illustrates this for $N = 2$, so (20) becomes $-\lambda(\beta_i^1 - \beta_i^2)I_1$. If sector 1 is locally

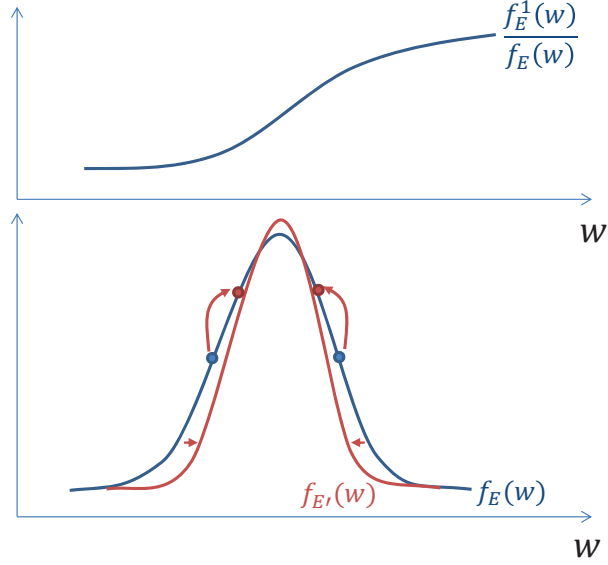


Figure 3: Incentive constraint effects

the high-wage sector and an increase in E_i reduces the relative return to activity 1, so $\beta_i^1 - \beta_i^2 < 0$, then it locally compresses the wage distribution. This is because higher wage individuals on average have a greater effort intensity in activity 1 and thus see their wage fall relative to lower wage, activity 2 intensive individuals. This moves wages closer together locally and is welfare improving. The total effect in (20) integrates over all these local wage compression effects.

3.3.3 Resource Constraint Effects

Because of the subshift 2 variation in schedules that undoes the average change in c and l from the marginal increase in E_i at each w , there are no net effects from (iii) here. We are therefore only left with the direct wage shift effect (ii). At each w , the average wage change (using (17)) is $w\delta_E^i(w)$, so the effect on the resource constraint is

$$\lambda \int_{\underline{w}_E}^{\bar{w}_E} \delta_E^i(w) w l(w) f_E(w) dw = \lambda \sum_{j=1}^N \beta_i^j(E) \int_{\underline{w}_E}^{\bar{w}_E} y(w) f_E^j(w) dw. \quad (22)$$

It is useful to write this in terms of the Pigouvian taxes $t_p^i(E)$, $i = 1, \dots, N$, defined by $r_i(E) - t_p^i(E) \equiv \partial Y(E) / \partial E_i$, i.e., as the tax on equivalent effort in sector i that fills the wedge between the private and social returns to i -sector effort (the corresponding tax on income in sector i is $\tau_p^i(E) = t_p^i(E) / r_i(E)$). Then $t_p^i(E)$ can be expressed as an output

weighted sum of the corrections for the externalities from E_i :

$$t_p^i(E) = - \sum_{j=1}^N \beta_i^j(E) Y^j(E). \quad (23)$$

In particular, if effort in activity i raises the returns to this and other activities, it generates positive externalities, so the Pigouvian tax is negative, and vice versa. Using this in (22) yields a resource constraint effect of simply $-\lambda t_p^i(E)$. Hence, a marginal increase in E_i increases (decreases) welfare through this effect if it generates positive (negative) externalities.

3.3.4 Consistency Constraint Effects

Consider the effects of a marginal increase in E_i on consistency constraint j . First, there is the direct effect (i), which is simply $\lambda \zeta_j \delta_{ij}$, where δ_{ij} is the Kronecker δ . Second, there are various effects on the right-hand side. For these, it is useful to rewrite consistency constraint j following (10) as $E_j = \int_{\Theta} \theta_j e_j(\theta) dF(\theta)$ and to note that

$$\theta_j e_j(\theta) = l(w_E(\theta)) \frac{\theta_j e_j(\theta)}{m(e(\theta))} = l(w_E(\theta)) \frac{\theta_j \frac{e_j(\theta)}{e_N(\theta)}}{m\left(\frac{e_1(\theta)}{e_N(\theta)}, \dots, \frac{e_{N-1}(\theta)}{e_N(\theta)}, 1\right)}$$

by linear homogeneity of m and the fact that $l = m(e)$. For the same reason and by (4), the effort ratios $\zeta_j \equiv e_j/e_N$ only depend on the vector of *relative returns*

$$x_E(\phi) \equiv \left(\phi_1 \frac{r_1(E)}{r_N(E)}, \dots, \phi_{N-1} \frac{r_{N-1}(E)}{r_N(E)}, 1 \right).$$

That is, $\zeta \equiv (\zeta_1, \dots, \zeta_{N-1}, 1)$ is homogenous of degree zero in the return vector $(\theta_1 r_1(E), \dots, \theta_N r_N(E))$. Using this, the effective effort integrated over on the right-hand side of consistency constraint j is

$$\theta_j e_j(\theta) = l(w) \theta_j \Omega_j(\zeta(x_E(\phi))) \quad \text{with} \quad \Omega_j(\zeta(x_E(\phi))) \equiv \frac{\zeta_j(x_E(\phi))}{m(\zeta(x_E(\phi)))}. \quad (24)$$

This reveals that there are two distinct effects here: first, the change in the overall level of effort $l(\cdot)$ for each individual (which is part of effect (iii)), holding constant the cross-sectoral allocation of efforts, and second, the re-allocation of effort across sectors due to the change in the relative returns x_E caused by the increase in E_i (effect (iv)).

Overall effort re-allocation effect. As for the former, the change in $l(\cdot)$ for individuals of wage w and type ϕ from subshift 1 is $w l'(w) \sum_{k=1}^N q_E^k(\phi) \beta_i^k(E)$, and using $\theta_j \Omega_j =$

$\theta_j e_j / l = w q^j / r_j$, the effect on (24) is

$$\frac{w q_E^j(\phi)}{r_j(E)} l'(w) w \sum_{k=1}^N q_E^k(\phi) \beta_i^k(E).$$

Averaging over the set $\{\theta | w_E(\theta) = w\}$ of all wage w individuals gives

$$\frac{1}{r_j(E)} \sum_{k=1}^N \beta_i^k(E) w^2 l'(w) \mathbb{E} \left[q_E^j(\phi) q_E^k(\phi) \mid w \right].$$

The average change in $l(\cdot)$ at w induced by the change in the l -schedule in subshift 2 is $-w l'(w) \sum_{k=1}^N \mathbb{E} [q_E^k(\phi) \mid w] \beta_i^k(E)$, and so the average change in sector j equivalent effort in (24) is

$$-\frac{1}{r_j(E)} \sum_{k=1}^N \beta_i^k(E) w^2 l'(w) \mathbb{E} \left[q_E^j(\phi) \mid w \right] \mathbb{E} \left[q_E^k(\phi) \mid w \right].$$

Integrating over all wages gives a total effect on consistency constraint j of

$$-\lambda \xi_j \sum_{k=1}^N \beta_i^k(E) C_{kj}(E), \quad (25)$$

where
$$C_{kj}(E) \equiv \frac{1}{r_j(E)} \int_{\bar{w}_E} w^2 l'(w) \text{Cov} \left(q_E^j, q_E^k \mid w \right) f_E(w) dw \quad (26)$$

with $\text{Cov} (q^j, q^k) = \mathbb{E} [q^j q^k] - \mathbb{E} [q^j] \mathbb{E} [q^k]$. The intuition is tightly linked to our variation: The schedule change in $l(w)$ is constructed to zero out the average change in effort at any given w , across all activities. If the j -sector income share q^j were uncorrelated with this effort change at any given w , then j -sector effort would also remain unchanged. If it were positively correlated, however, then it would increase, and vice versa. In particular, if a marginal increase in E_i increases the returns to activities k in which individuals have a high income share who also earn a lot of their income in activity j , then individuals with a high q^j see their wage increase more than proportionally and therefore move up along the $l(w)$ -schedule relative to others. Hence, if $l'(w) \geq 0$, this variation effectively re-allocates effort towards activity j , thus increasing the right-hand side of consistency constraint j .

Figure 4 illustrates this for $N = 2$, so by (25) the effect on consistency constraint 1 becomes $-\lambda \xi_1 (\beta_i^1 - \beta_i^2) C_{11}$ and the effect on consistency constraint 2 is $\lambda \xi_2 (\beta_i^1 - \beta_i^2) C_{22}$, where we used $q_E^2 = 1 - q_E^1$. Suppose an increase in E_i increases the relative return to activity 1, as drawn in the figure, so $\beta_i^1 - \beta_i^2 > 0$. This has different effects on individuals who earned the same wage \bar{w} before the change (and thus were located on the same iso-wage curve in θ -space in the left panel) but did so using different intensities. Those

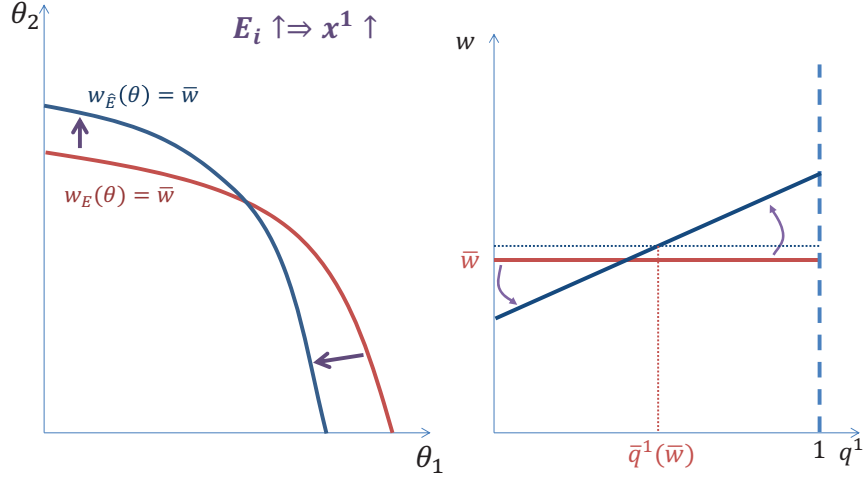


Figure 4: Overall effort re-allocation effect

with a high intensity q^1 in activity 1 (because of a high θ_1/θ_2 -ratio) see their wage increase whereas the others see it fall. If $l'(w) \geq 0$, so $C_{11}, C_{22} \geq 0$, the former group will increase their effort and the latter reduce it, thus leading to a flow of effort out of sector 2 and into 1. The figure also makes clear that this effect relies on there being variation in intensities at a given wage: if every wage- \bar{w} earner had the same (average) intensity $\bar{q}^1(\bar{w})$ and so $\text{Var}(q^1|\bar{w}) = 0$, then the rotation shown in the right panel would be without consequences, since $C_{11}, C_{22} = 0$.

Sectoral shift effect. Second, the effect of E_i through the change in the vector of effort ratios ζ on (24) is,⁹

$$l(w)\theta_j \sum_{k=1}^N \sum_{l=1}^N \frac{\partial \Omega_j(\zeta(x_E(\phi)))}{\partial \zeta_l} \frac{\partial \zeta_l(x_E(\phi))}{\partial (\theta_k r_k(E))} \frac{\partial \theta_k r_k(E)}{\partial E_i}. \quad (27)$$

We can rewrite this using $q^j = r_j \theta_j \Omega_j / w$ and hence $q_E^j(\phi) = Z_j(x_E(\phi)) \Omega_j(\zeta(x_E(\phi)))$ with

$$Z_j(x_E(\phi)) \equiv \frac{r_j(E)\theta_j}{w} = x_E^j(\phi) \min_{p \in \Delta^{N-1}} m \left(\frac{p_1}{x_E^1(\phi)}, \dots, \frac{p_{N-1}}{x_E^{N-1}(\phi)}, p_N \right), \quad (28)$$

where we used (3) and homogeneity of degree one of m . This makes explicit that the income share q^j is also only a function of the relative returns x , and in fact is a product of two terms, one of which, Z_j , directly depends on x and the other of which, Ω_j , depends

⁹Our assumptions on m ensure that $\zeta_l(x_E(\phi))$ is continuous and differentiable almost everywhere—with potential non-differentiabilities only at the borders of corners where $\zeta_l(x_E(\phi)) = 0$.

on x only through the vector of effort ratios ζ . It is instructive to rewrite (27) by defining

$$Q_k^j(x_E(\phi)) \equiv Z_j(x_E(\phi)) \sum_{l=1}^N \frac{\partial \Omega_j(\zeta(x_E(\phi)))}{\partial \zeta_l} \frac{\partial \zeta_l(x_E(\phi))}{\partial (r_k(E)\theta_k)} \theta_N r_N(E). \quad (29)$$

Intuitively, there are two components to the effect of a change in the return to sector k relative to N (holding all other returns constant relative to sector N) on the sectoral income share q^j : the *mechanical* effect of changing returns (through Z_j), holding constant all efforts, and the indirect *sectoral shift* effect (through Ω_j) due to the reallocation of effort across sectors. Q_k^j measures only the latter component, i.e., the substitution effect on Q^j that results from the change in the sectoral effort ratios ζ in response to a change in the sector- k returns $r_k(E)\theta_k$, but holding relative returns x fixed otherwise. Substituting this in (27), the effect becomes

$$l(w)\theta_j \sum_{k=1}^N \frac{w}{r_j(E)\theta_j} Q_k^j(x_E(\phi)) \frac{\partial r_k(E)\theta_k}{\partial E_i} \frac{1}{r_N(E)\theta_N} = \frac{y(w)}{r_j(E)} \sum_{k=1}^N \beta_i^k(E) Q_k^j(x_E(\phi)) x_E^k(\phi)$$

since $\partial(r_k(E)\theta_k)/\partial E_i = \beta_i^k(E)r_k(E)\theta_k$. Integrating over all wages and all ϕ gives a total effect on consistency constraint j of

$$- \lambda \tilde{\zeta}_j \sum_k \beta_i^k(E) S_{kj}(E) \quad (30)$$

with
$$S_{kj}(E) \equiv \frac{1}{r_j(E)} \int_{\bar{w}_E} y(w) \int_{\Phi} Q_k^j(x_E(\phi)) x_E^k(\phi) dG_E(w, \phi). \quad (31)$$

As a result, if a marginal increase in E_i increases the returns to activities k (i.e. $\beta_i^k(E) > 0$) for which $Q_k^j > 0$ (so that an increase in $r_k\theta_k$ increases the income share earned in activity j through a re-allocation of the effort ratios towards e_j), then it induces a shift of effort into sector j and thus increases the right-hand side of the j -th consistency constraint. This is again illustrated in Figure 5 for $N = 2$. If an increase in E_i increases the relative return r_1/r_2 to activity 1, individuals will shift their effort allocation e_1/e_2 towards activity 1, thus reinforcing the overall effort reallocation effect illustrated in Figure 4.

Adding up. Per the preceding discussion, the C_{kj} and S_{kj} effects can both be interpreted as across sector re-allocations. Formally, as the following lemma shows, the shifts of *incomes* across sectors induced by those two effects have to sum to zero across all j —i.e., the r_j -weighted rows sum to zero. The lemma also establishes the more mechanical facts that the columns of C_{kj} and S_{kj} sum to zero.

Lemma 3. (i) $\sum_{j=1}^N r_j(E)C_{kj}(E) = \sum_{j=1}^N r_j(E)S_{kj}(E) = 0$ for all $k = 1, \dots, N$.
(ii) $\sum_{k=1}^N C_{kj}(E) = \sum_{k=1}^N S_{kj}(E) = 0$ for all $j = 1, \dots, N$.

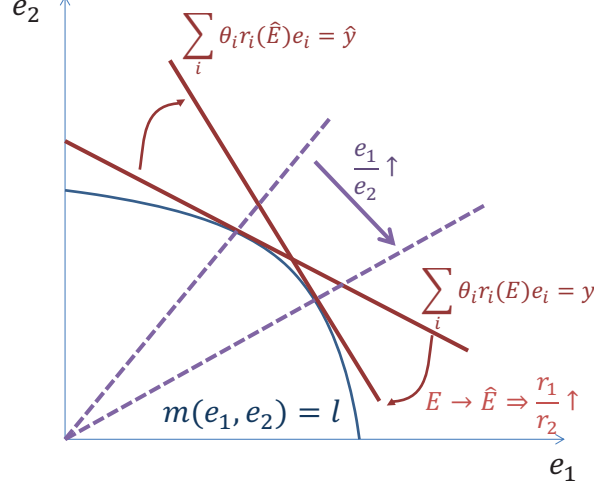


Figure 5: Sectoral shift effect

The intuition for part (i) of Lemma 3 hinges on the fact that the $r_j(E)$ -weighted sum of the right-hand sides of the N constraints in (8) is $\int_{\bar{w}_E}^{\bar{w}_E} y(w) f_E(w) dw$, and the sectoral composition of income at wage w is irrelevant for this sum. Per (29) and the subsequent discussion, the changes $\{Q_k^j\}_{j=1, \dots, N}$ in $\{S_{kj}\}_{k=1, \dots, N}$ reflect changes in the income shares q^j , and thus do not affect $\int_{\bar{w}_E}^{\bar{w}_E} y(w) f_E(w) dw$. Similarly, subshift 2 ensures, by construction, that $l(w)$ and hence (for effect (iii)) $wl(w) = y(w)$ is unchanged on average at each w . Hence, both effects reflect income shifts across activities in response to return changes, which have to add up to zero.

The converse adding-up property in part (ii) of the lemma is simply due to the fact that proportional changes in *all* returns (i) do not affect individuals' cross-sectoral effort allocation since m is linear homogeneous, so there are no sectoral shift effects, and (ii) cause equi-proportional changes in the wages of all types θ in subshift 1, and hence no cross-sectoral re-allocation of overall effort $l(w)$ at any wage. A direct consequence of this is that (25) and (30) are non-zero only if a change in E_i affects *relative* returns.

3.3.5 Putting Them Together

To find the total welfare effect of a marginal change in E_i , we combine (18), (20), (22), with (25), (30) and the direct effect $\xi_j \delta_{ij}$ for all consistency constraints j . Moreover, because of the adding-up property in Lemma 3 (ii) and the fact that $\sum_j I_j = \sum_j R_j = 0$, only relative return changes matter for the effects (18), (20), (25) and (30), so we can equivalently write (18) as $-\lambda \sum_j (\beta_i^j - \beta_i^N) R_j$ and analogously for the others. Defining $\Delta \beta_i^j(E) \equiv \beta_i^j(E) - \beta_i^N(E)$ as the semi-elasticity of the returns $(\partial x_E^j(\phi) / \partial E_i) / x_E^j(\phi)$ establishes the following lemma, which summarizes the results from this subsection:

Lemma 4. *At any Pareto optimum with $E_i > 0$, the welfare effect of a marginal change in E_i is*

$$\frac{\partial W(E)}{\partial E_i} = \lambda \left[\xi_i - t_p^i(E) - \sum_j \Delta\beta_i^j(E) \left(I_j(E) + R_j(E) + \sum_k \xi_k (C_{jk}(E) + S_{jk}(E)) \right) \right],$$

with $R_j(E)$, $I_j(E)$, $t_p^i(E)$, $C_{jk}(E)$ and $S_{jk}(E)$ respectively given by (19), (21), (23), (26) and (31).

This makes clear that, if $\Delta\beta_i^j = 0$ for all j , i.e., an increase in E_i has no effect on the vector of relative returns x , then $\xi_i = t_p^i(E)$ at the optimum. Any deviation of ξ_i from $t_p^i(E)$ is due to the relative return effects I , R , C and S .

3.4 Marginal Tax Rate Results and Outer Problem Dimensionality

Define $\Delta\beta$, C , and S , respectively, as the matrices with (i, j) th elements $\Delta\beta_i^j(E)$, $C_{ij}(E)$, and $S_{ij}(E)$. Define \vec{I} and \vec{R} as the column vectors with elements $I_i(E)$ and $R_i(E)$, respectively. Finally, define $\vec{\xi}$ and \vec{t}_p respectively as the column vectors with elements ξ_i and $t_p^i(E)$, and use \mathbf{I}_N to denote the $N \times N$ identity matrix. If the optimal allocation has $E_i > 0$ for all i ,¹⁰ the N of optimality conditions for E_i from Lemma 4 can be written as

$$(\mathbf{I}_N - \Delta\beta(C + S))\vec{\xi} = \vec{t}_p + \Delta\beta(\vec{I} + \vec{R}). \quad (32)$$

Because the returns $r_i(E)$ are endogenous, individuals may impose externalities when they exert effort to earn income in sector i . $t_p^i(E)$ is the tax on sector- i effective effort needed to align an individual's private and social returns to sector- i effort in the partial-equilibrium sense—i.e., holding fixed others' behavior (and, as discussed above, $\tau_p^i = t_p^i/r_i$ is the corresponding tax on sector- i income). If there were no general-equilibrium effects, we would expect the Pigouvian corrective tax on income earned by wage w individuals to be $\sum_i (f_E^i(w)/f_E(w)) \tau_p^i(E)$, i.e., a weighted average of the sector-specific Pigouvian corrections, with weights reflecting the share of income earned in the various sectors. Intuitively, an increase in the marginal income tax rate at a given income does not directly affect any individual's optimal sectoral income shares q^i , so a marginal tax distorts average sectoral incomes $y_i(w)$ earned by individuals at any w in proportion to the income share $f_E^i(w)/f_E(w)$ at w .

Of course, there will typically be indirect, general equilibrium effects as well: changes in E change returns $r_i(E)$ and hence individuals' optimal cross-sectoral effort allocations. The *actual* externality correction term $\sum_i (f_E^i(w)/f_E(w)) (\xi_i/r_i)$ from (12) includes these

¹⁰If $E_i = 0$, the following analysis would apply to the subset of non-zero E_i .

general equilibrium effects. We can use the system (32) to compare the general- and partial-equilibrium corrections. They would obviously coincide if $\tau_p^i(E) = \xi_i/r_i$, or equivalently $t_p^i(E) = \xi_i$, for each sector i , as would be the case if $\Delta\beta$ were identically zero and there were no relative effects of *any* change in E .

More generally, the general- and partial-equilibrium corrections at wage w coincide whenever the proportional change in income $\Delta Y^i \propto f_E^i(w)/f_E(w)$ induced by an additional marginal income tax at wage w implies an aggregate effort change $\Delta E_i = \Delta Y^i/r_i$ in a *direction* in which there are no relative return effects. Formally, let \vec{n} denote the column vector with i^{th} element $n_i = (f_E^i(w)/f_E(w))(1/r_i(E))$ and \vec{n}' its transpose. If $\vec{n}'\Delta\beta = 0$, so that there are no relative wage effects in the direction \vec{n}' , then left-multiplying (32) yields $\vec{n}'\vec{\xi} = \vec{n}'\vec{t}_p$, i.e.,

$$\sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i(E)} = \sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \tau_p^i(E)$$

and the general equilibrium correction coincides with the partial equilibrium correction.¹¹

Intuitively, any wedge between the partial- and general-equilibrium corrections is attributable to relative return effects: if a marginal tax at wage w does not cause any change in relative returns, then imposing it will change the aggregate efforts E , but it will not change any type's optimal cross-sectoral effort allocation, and hence will have no indirect feedback effects on E . When there are relative return effects, and $\vec{n}'\Delta\beta \neq 0$, then the optimal correction, per (12), will in general diverge from the partial-equilibrium correction.

We can also ask when, additionally, the optimal correction in (12) is zero—i.e., when $\sum_i (f_E^i(w)/f_E(w)) (\xi_i/r_i) = \sum_i (f_E^i(w)/f_E(w)) \tau_p^i(E) = 0$, so that the marginal tax rate formula (12) is the same as in a standard one-dimensional Mirrlees model. The following result provides a simple characterization.

Proposition 2. *Suppose $Y(E) > 0$. Then \vec{n}' is a direction of both no relative return effects and no externalities, i.e., $\vec{n}'\Delta\beta = 0$ and $\vec{n}'\vec{t}_p = 0$, if and only if it is a left-nullvector of β : $\vec{n}'\beta = 0$.*

Let $N - K$ denote the rank of the $\mathbb{R}^N \rightarrow \mathbb{R}^N$ mapping $r(E) = (r_1(E), \dots, r_N(E))'$ and hence of the matrix β . Since the return vector $r(E)$ is a sufficient statistic for individual behavior, conditional on a given tax code (equivalently, an $l(w)$ - and $V(w)$ -schedule),¹² one might hope to reduce the dimensionality of the outer problem when $K > 0$ —i.e., whenever, by Proposition 2, there exist directions \vec{n}' in which there are both no externalities and no relative return effects.

¹¹It is worth noting that $\Delta\beta$ is singular (its last column is identically zero), so it has a non-empty left-nullspace. There are no relative return effects at w precisely when \vec{n}' lies in this left-nullspace.

¹²So that $r(E)$ is also sufficient for describing the variation, with E , of the tax code that leads to (32).

This is the case, for example, in the framework of Rothschild and Scheuer (2013b), where $N = 2$ and $Y(E)$ has constant returns to scale with $r_i(E) = Y_i(E)$, so that private returns equal marginal products. Since the latter are homogeneous of degree zero, they are only a function of $\rho \equiv E_1/E_2$, and it is easy to verify that the second row of β is just $-\rho$ times the first row. In other words, β has rank $N - K = 1$ for all E , and, as shown by Rothschild and Scheuer (2013b), the outer problem can be written in terms of the *single* variable ρ and with a single consistency constraint

$$\rho = \frac{\int_{\underline{w}_E}^{\bar{w}_E} w l(w) dF_\rho^1(w) / r_1(\rho)}{\int_{\underline{w}_E}^{\bar{w}_E} w l(w) dF_\rho^2(w) / r_2(\rho)}.$$

Similar reductions in dimensionality can occur for $N > 2$. Suppose, for instance, $N = 3$ and $\beta_2^j(E) = a\beta_1^j(E)$ and $\beta_3^j(E) = b\beta_1^j(E)$ for all j , where a and b are constants. Here, E_1 , E_2 and E_3 have effects on the returns r_j that only differ in their magnitude or sign (in percentage terms). Then there is a two-dimensional plane with directions of no relative return effects and no externalities spanned by the vectors $(-a, 1, 0)$ and $(-b, 0, 1)$. The vector orthogonal to both is $(1, a, b)$, so $\tilde{E}_1 = E_1 + aE_2 + bE_3$ is a sufficient statistic for the return vector $r(E)$. The outer problem can again be written with a single consistency constraint, namely a weighted average of the three consistency constraints in (8).

In fact, the following proposition shows that the dimensionality of the outer problem can be reduced with a proper choice of coordinates whenever the rank of β is less than N .

Proposition 3. *Suppose that β has rank $N - K$ in some open neighborhood of the optimum E^* . Then there exists an open neighborhood $U \in \mathbb{R}^N$ on which the Pareto problem can be written as a function of the schedules $l(w)$, $V(w)$, and some $\rho \in \mathbb{R}^{N-K}$ and with $N - K$ consistency constraints, one for each component of ρ .*

Whenever we can reduce dimensionality through a change in coordinates, we can, of course, also reformulate the marginal tax rate formula from the inner problem (12) in terms of this new basis. As before, let \vec{n} be the vector with elements $f_E^i(w) / (f_E(w)r_i(E))$, $\vec{\zeta}$ the vector with elements ζ_i , where ζ_i are the multipliers in the original E coordinate system, and \mathbf{L} as the matrix with elements $L_{ij} = \partial G_i(E^*) / \partial E_j$ (at an optimum E^*), where G is the coordinate transformation defined in the proof of Proposition 3. Then the correction term in (12) in terms of E , $\sum_{i=1}^N (f_E^i(w) / f_E(w)) (\zeta_i / r_i(E))$, can be replaced in the new coordinates with $\vec{n}' \mathbf{L} \vec{\zeta}$, where $\vec{\zeta}$ is the vector whose first $N - K$ elements are the multipliers of the consistency constraints in the new coordinate system and last K elements are zero.

Finally, for the system of optimality conditions (32) to be fully informative about the optimal vector $\vec{\zeta}$, the matrix $\mathbf{A} \equiv \mathbf{I}_N - \Delta\beta(\mathbf{C} + \mathbf{S})$ needs to be invertible and hence non-

singular at the optimum, which we assume in the following. If \mathbf{A} had less than full rank, there would be multiple solutions for $\vec{\zeta}$ in (32), and so the outer problem variation we used to obtain this system would not fully identify $\vec{\zeta}$. This full-rank assumption implies that all eigenvalues of \mathbf{A} must be nonzero. In fact, defining \vec{r} as the column vector of sectoral returns $r_i(E)$, the adding up property of the sectoral shift matrices \mathbf{C} and \mathbf{S} in Lemma 3 (i) immediately implies that $\mathbf{A}\vec{r} = \vec{r}$, so that \vec{r} is always an eigenvector of \mathbf{A} with associated eigenvalue 1.

For the analysis of some of the examples later on, we will assume that all other eigenvalues of \mathbf{A} are not only nonzero, but strictly positive as well. In Appendix B, we show that this assumption corresponds to a notion of stability of the fixed point for E at the optimum, which is closely related to the variation underlying (32).

4 Two Sectors

If $N = 2$, we can use the system of optimality conditions (32) to solve for $\vec{\zeta}$ explicitly. Note first that, in this case, $q_E^2(\phi) = 1 - q_E^1(\phi)$ and so $\text{Cov}(q_E^1, q_E^2|w) = -\text{Var}(q_E^1|w)$. Moreover, by the proof of Lemma 3, $Q_1^2(x_E(\phi)) = -Q_1^1(x_E(\phi))$, so that we can write

$$\Delta\beta(\mathbf{C} + \mathbf{S}) = \begin{pmatrix} \Delta\beta_1^1(E)/r_1(E) & -\Delta\beta_1^1(E)/r_2(E) \\ \Delta\beta_2^1(E)/r_1(E) & -\Delta\beta_2^1(E)/r_2(E) \end{pmatrix} (\mathbf{C}(E) + \mathbf{S}(E))$$

where

$$\mathbf{C}(E) \equiv r_1(E)\mathbf{C}_{11}(E) = \int_{\underline{w}_E}^{\bar{w}_E} w^2 l'(w) \text{Var}(q_E^1|w) f_E(w) dw \quad (33)$$

and

$$\mathbf{S}(E) \equiv r_1(E)\mathbf{S}_{11}(E) = \int_{\underline{w}_E}^{\bar{w}_E} y(w) \int_{\Phi} Q_1^1(x_E^1(\phi)) x_E^1(\phi) dG_E(w, \phi). \quad (34)$$

Since $Q_1^1 > 0$ (the substitution effect leads to a re-allocation of effort towards activity 1 when the relative return to this activity increases), we have $S(E) \geq 0$, and also $C(E) \geq 0$ if $l(w)$ is increasing in w at the optimum.

Since $(\Delta\beta_2^1(E), -\Delta\beta_1^1(E))$ is always a left-nullvector of $\Delta\beta$ and therefore a direction of no relative return effects when $N = 2$, we can interpret the orthogonal direction $(\Delta\beta_1^1(E), \Delta\beta_2^1(E))$ as the direction of *maximal* relative return effects. Let $\gamma_2(E)$ be the second, non-unit eigenvalue of the matrix $\mathbf{A} = \mathbf{I}_2 - \Delta\beta(\mathbf{C} + \mathbf{S})$, i.e.,

$$\gamma_2(E) = 1 + \left(\frac{\Delta\beta_2^1(E)}{r_2(E)} - \frac{\Delta\beta_1^1(E)}{r_1(E)} \right) (\mathbf{C}(E) + \mathbf{S}(E)), \quad (35)$$

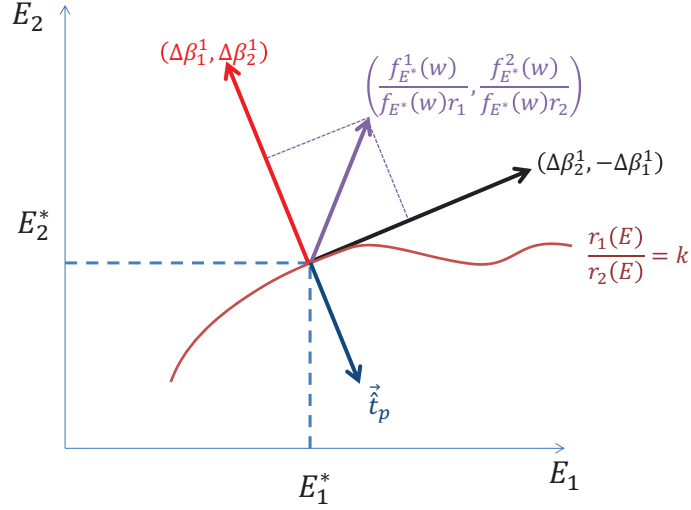


Figure 6: Adjustment factor and directions of no and maximal relative return effects

which is associated with the eigenvector $(\Delta\beta_1^1(E), \Delta\beta_2^1(E))'$. If the optimum involves a stable fixed point for E in the sense of our discussion in Appendix B, then $\gamma_2(E) > 0$.

Using this, solving system (32) yields:

Lemma 5. *At any Pareto optimum with $N = 2$,*

$$\vec{\zeta} = \vec{t}_p + \begin{pmatrix} \Delta\beta_1^1(E) \\ \Delta\beta_2^1(E) \end{pmatrix} \frac{I_1(E) + R_1(E) + (\tau_p^1(E) - \tau_p^2(E)) (C(E) + S(E))}{\gamma_2(E)}. \quad (36)$$

The system (36) makes it easy to interpret the corrective term in the marginal tax rate formula (12). As before, we obtain $\zeta_i = t_p^i(E)$ if $\Delta\beta_i^1(E) = 0$, $i = 1, 2$, so that a change in E_i has no relative return effects at the optimum. More generally, if the vector \vec{n} with elements $f_E^i(w)/(r_i(E)f_E(w))$ is parallel to the direction of no relative return effects $(\Delta\beta_2^1(E), -\Delta\beta_1^1(E))$, then the marginal tax rate formula (12) coincides with the weighted sum of the partial equilibrium Pigouvian corrections, as discussed for the case of general N in section 3.4, so that $\vec{n}'\vec{\zeta} = \vec{n}'\vec{t}_p$. For any other \vec{n} , the correction term $\vec{n}'\vec{\zeta}$ will diverge from the Pigouvian correction $\vec{n}'\vec{t}_p$, with the magnitude of this divergence determined by the magnitude of the second term in (36) and the angle between \vec{n} and the direction of no relative return effects $(\Delta\beta_2^1(E), -\Delta\beta_1^1(E))$. This is illustrated in Figure 6, which shows an iso-relative return curve in (E_1, E_2) -space as well as the (tangent) direction of no relative return effects starting from a Pareto optimum (E_1^*, E_2^*) , the (perpendicular) direction of maximal relative return effects, and the projection of the vector \vec{n} on the latter.

By Propositions 2 and 3, the outer problem can be reduced, via an appropriate change of variables, to a one-dimensional problem whenever the direction of no relative return

effects is also a direction of no externalities or, equivalently, whenever \vec{t}_p is parallel to the direction of maximal relative return effects: $\vec{t}_p = x(\Delta\beta_1^1, \Delta\beta_2^1)$ for some x , as is the case in Figure 6. Clearly, this is trivially the case when there are no externalities (so that $x = 0$, see Rothschild and Scheuer, 2013b) or only one activity affects returns (so that $\Delta\beta_2^1 = t_p^2 = 0$, see Rothschild and Scheuer, 2013a), as we will discuss in more detail below.

In the remainder of the paper, we will use formula (36) in order to sign the deviation of ξ_i from t_p^i in various applications of our general model, and to explore its implications for the shape of the optimal non-linear income tax schedule.

5 Applications

5.1 No Externalities

Suppose $Y(E)$ is homothetic, and let us consider, to begin, the externality-free case where private returns coincide with marginal products: $r_i(E) = Y_i(E)$ for all i . The following lemma shows that, in this case, $Y(E)$ must exhibit constant returns to scale:

Lemma 6. *If $Y(E)$ is homothetic and there are no externalities then $Y(E)$ is linear homogeneous.*

In Rothschild and Scheuer (2013b), we have considered the special case with $N = 2$. The general framework developed in Section 3 can be used to investigate the novel effects that can arise when sectoral choice is along more than one margin. The simplest way to shed light on this is to add a third, linear sector, leading to the production function $Y(E) = \hat{Y}(E_1, E_2) + E_3$, where \hat{Y} exhibits constant returns to scale and $r_i(E) = Y_i(E)$ for $i = 1, 2, 3$. This is particularly tractable as E_3 has no effects on any returns and $\rho \equiv E_1/E_2$ remains a sufficient statistic for the wage distribution. The general system of optimality conditions (32) can be solved to obtain the following modified marginal tax rate adjustment factor.

Proposition 4. *If $N = 3$, $Y(E) = \hat{Y}(E_1, E_2) + E_3$, \hat{Y} is homothetic and there are no externalities, then the numerator in the marginal tax rate formula (12) is*

$$1 - \sum_{i=1}^3 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i(E)} = 1 + \frac{f_E^1(w) + f_E^2(w)}{f_E(w)} \left(\frac{f_E^1(w)}{f_E^1(w) + f_E^2(w)} - \hat{\alpha}(\rho) \right) \hat{\xi} \quad (37)$$

with

$$\hat{\xi} = -\frac{1}{r_1(1 - \hat{\alpha})} \frac{\beta_1^1(I_1 + R_1) + \beta_1^2(I_2 + R_2)}{1 - \beta_1^1(C_{11} + S_{11} - \rho(C_{12} + S_{12})) - \beta_1^2(C_{21} + S_{21} - \rho(C_{22} + S_{22}))}, \quad (38)$$

where $\hat{\alpha}(\rho) \equiv Y^1/\hat{Y}$ is the sector 1 income share relative the income of sectors 1 and 2 together.

In the two-sector case where E_3 and $f_E^3(w)$ vanish, Lemma 3 implies $C_{21} = -C_{11}$, $r_1C_{12} = -r_2C_{11}$, and $r_2C_{22} = r_1C_{11}$, and analogously for the S_{ij} terms. Hence, knowing any one of the four flow terms $C_{ij} + S_{ij}$ determines the other three. Similarly, $I_2 + R_2 = -(I_1 + R_1)$. Denoting by $\sigma(\rho)$ the substitution elasticity of $Y(E)$ and by $\alpha(\rho) \equiv Y^1(E)/Y(E)$ the aggregate income share of sector 1, the adjustment factor then collapses to the formula from Rothschild and Scheuer (2013b):

Corollary 1. *If $N = 2$, $Y(E)$ is homothetic, and there are no externalities, then the numerator in the marginal tax rate formula (12) is*

$$1 - \sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\zeta_i}{r_i(E)} = 1 + \left(\frac{f_E^1(w)}{f_E(w)} - \alpha(\rho) \right) \bar{\xi} \quad \text{with} \quad \bar{\xi} \equiv \frac{(I_1 + R_1) / \sigma}{\alpha(1 - \alpha)Y + (C + S) / \sigma}, \quad (39)$$

where $C = r_1C_{11}$ and $S = r_1S_{11}$ by (33) and (34).

This makes clear that the adjustment factor vanishes as $\sigma(\rho) \rightarrow \infty$ and technology becomes linear. Otherwise, suppose sector 1 is the high-wage sector and redistributive motives at least weakly favor the low-wage sector 2, so that $I_1(E) > 0$ and $R_1(E) \geq 0$. Then $\bar{\xi} > 0$, so marginal keep shares are scaled up compared to the standard formula in parts of the wage distribution where sector 1 is prevalent and hence its local income share $f_E^1(w)/f_E(w)$ exceeds its aggregate income share $\alpha(\rho)$, and scaled down otherwise. In other words, marginal tax rates are scaled down for high wages and scaled up for low wages, making the tax schedule less progressive than in a standard Mirrlees model; for example, the top marginal tax rate is $T'(y(w)) = (\alpha - f_E^1(w)/f_E(w)) \bar{\xi} < 0$.

The intuition is that the optimal income tax makes use of general equilibrium effects to indirectly redistribute from high to low wage earners, introducing a regressive force when the sectors are complementary: lowering taxes at wages where activity 1 is prevalent increases ρ and therefore increases the returns to the low-wage activity 2.¹³ On the other hand, the sectoral reallocation effects $C + S > 0$ lower $\bar{\xi}$ and hence blunt the magnitude of the regressive adjustment to the tax schedule relative to a model with fixed sectoral efforts. Intuitively, a regressive tax beneficially redistributes via an increase in the relative returns to the low-income sector 2; but in so doing, it leads individuals to shift effort into sector 2, which counteracts the original relative return effect. The optimal tax schedule is therefore more regressive than in a Mirrlees (1971) model with fixed wages, but less regressive than in an endogenous wage model with fixed occupations like Stiglitz (1982).

¹³Note that constant returns to scale and concavity of Y imply $\sigma \geq 0$. Also, the exact same results obtain if sector 2 is the high-wage sector and sector 1 the low-wage sector, and redistributive motives again favor the low-wage sector.

The adjustment vanishes at wage levels w where $f_E^1(w)/f_E(w) = \alpha$, so that the local and aggregate income shares coincide. This is because, there, $\vec{n}' = (f_E^1(w)/(f_E(w)Y_1), f_E^2(w)/(f_E(w)Y_2))$ reduces to $(E_1/Y, E_2/Y)$ and therefore points in the direction of no relative return effects $(\rho, 1)$, as discussed in section 3.4. Moreover, any direction here is trivially a direction of no externalities, any nonzero adjustment term in (12) is exclusively due to relative return effects.

When the third sector is active, the adjustment factor in (37) similarly vanishes whenever $f_E^1(w)/(f_E^1(w) + f_E^2(w)) = \hat{\alpha}$. However, it is now scaled down by the local share $(f_E^1(w) + f_E^2(w))/f_E(w)$ of income earned in sectors 1 and 2, reflecting the fact that the relative return effects only operate through this fraction of the population. In addition, the term $\hat{\zeta}$ in (38) will generally diverge from the term $\bar{\zeta}$ in (39). This is because, although the *formula* for $\hat{\zeta}$ in (38) is the same in a two- and a three-sector model, the adding up properties in Lemma 3 that pin down the relationship between the four $C_{ij} + S_{ij}$, $i, j = 1, 2$ terms in a two-sector model do *not* pin them down when there is a third sector. Nevertheless, it is instructive to use Lemma 3 to re-write (38) as

$$\hat{\zeta} \equiv \frac{(I_1 + R_1) / \hat{\sigma} + E_1 \beta_1^2 (I_3 + R_3)}{\hat{\alpha} (1 - \hat{\alpha}) \hat{Y} + \frac{\hat{C}_{11} + \hat{S}_{11}}{\hat{\sigma}} + \hat{\alpha} \frac{\hat{C}_{13} + \hat{S}_{13}}{\hat{\sigma}} + \hat{\alpha} (1 - \hat{\alpha}) E_1 \beta_1^2 \left[\frac{\hat{C}_{31} + \hat{S}_{31}}{\hat{\alpha}} - \frac{\hat{C}_{32} + \hat{S}_{32}}{1 - \hat{\alpha}} \right]}, \quad (40)$$

where $\hat{\sigma}(\rho)$ is the substitution elasticity of \hat{Y} and $\hat{C}_{ij} \equiv r_j C_{ij}$, $\hat{S}_{ij} \equiv r_j S_{ij}$. Comparing with (39) reveals an extra term in the numerator and two extra terms in the denominator.

The first extra term in the denominator, namely $\hat{\alpha}(\hat{C}_{13} + \hat{S}_{13})/\hat{\sigma}$, arises formally from the row adding up constraint in Lemma 3: with two sectors, any outflow $-(\hat{C}_{11} + \hat{S}_{11})$ of sector 1 earnings caused by the r_1 -decrease associated with increased E_1 flows mechanically into sector 2; with a third sector, some will flow instead to sector 3 (whenever $\hat{C}_{13} + \hat{S}_{13} < 0$). Unlike flows into sector 2, sector 3 inflows do not decrease $\rho = E_1/E_2$, however, and therefore do not further blunt the desirable effects of an increase in E_1 . Hence, relative to the two-sector model, the presence of the third sector, reflected in this term, makes $\hat{\zeta}$ larger, and the optimal tax more regressive. Informally: the third sector moves the optimal tax closer to Stiglitz (1982) and away from Mirrlees (1971).

The second extra term in the denominator, whose sign only depends on the difference in square brackets because $\beta_1^2 > 0$, arises from property (ii) in Lemma 3. With two sectors, a proportional increase in both r_1 and r_2 induces no sectoral shifts and hence has no effect on ρ . However with three sectors, a proportional increase in r_1 and r_2 induces an income shift out of sector 3 and into sectors 1 and 2, which affect ρ insofar as the term in square brackets is non-zero: this occurs precisely when a change in r_3 leads to unequal percentage changes in sector 1 and 2 incomes (which, by Lemma 3 (ii), im-

plies the same for a proportional and simultaneous change in r_1 and r_2). For example, if $(\hat{C}_{31} + \hat{S}_{31})/\hat{\alpha} < (\hat{C}_{32} + \hat{S}_{32})/(1 - \hat{\alpha}) < 0$, sector 1 income changes by relatively more than sector 2 income when r_1 and r_2 change proportionally. This means that the effects of r_1 are effectively larger than the effects of r_2 . In our $\hat{C}_{11} + \hat{S}_{11}$ normalization in (40), the r_2 -effects of an increase in E_1 are therefore reduced, reinforcing the first extra effect.

Finally, assuming for simplicity $R_3 = 0$, the extra term in the numerator can be understood as follows. (Analyzing $R_3 \neq 0$ is straightforward.) Suppose that $f_E^1(w)/f_E(w)$ is increasing in w , so sector 1 is the high income sector. With two sectors, this mechanically implies that sector 2 must be the low income sector. The wage changes induced by an increase in E_1 thus beneficially redistribute, by reducing r_1 and increasing r_2 , from the high income sector 1 to the low income sector 2. But if there is a third sector and $f_E^3(w)/f_E(w)$ is decreasing, so that $I_3 < 0$, then sector 2 is less of a low income sector (in fact, $f_E^2(w)/f_E(w)$ could be increasing), so the wage changes induced by an increase in E_1 are distributionally less desirable. On the other hand, if $I_3 > 0$ so sector 3 is also a high-income sector, then sector 2 must be an even lower income sector than otherwise, so the indirect redistributive benefits of an increase in E_1 are magnified, and so is $\hat{\xi}$ and the regressive adjustment to the tax schedule.

5.2 Increasing or Decreasing Returns to Scale

Now let $Y(E)$ be any homothetic production function. Then it can be written as $Y(E) = h(\tilde{Y}(E))$, where $h(\tilde{Y})$ is some increasing function and $\tilde{Y}(E)$ is a constant returns to scale production function as in the preceding subsection. Let $N = 2$ and denote the substitution elasticity of $\tilde{Y}(E)$ as before by $\sigma(\rho)$ and the sector 1 income share by $\alpha(\rho) \equiv \tilde{Y}^1(E)/\tilde{Y}(E)$ with $\rho = E_1/E_2$. Suppose the total output Y is divided across sectors according to the \tilde{Y} -income shares, i.e. $Y^1(E) = \alpha(\rho)Y(E)$ and $Y^2(E) = (1 - \alpha(\rho))Y(E)$. Then by linear private returns,

$$r_1(E) = \alpha(\rho)Y(E)/E_1 \quad \text{and} \quad r_2(E) = (1 - \alpha(\rho))Y(E)/E_2. \quad (41)$$

Denoting the elasticity of h by $\varepsilon_h(E) \equiv h'(\tilde{Y}(E))\tilde{Y}(E)/Y(E)$, we have increasing returns to scale when $\varepsilon_h > 1$ and decreasing returns to scale when $\varepsilon_h < 1$. Lemma 5 then yields the following result:

Proposition 5. *If $N = 2$, $Y(E) = h(\tilde{Y}(E))$ with $\tilde{Y}(E)$ linear homogeneous, and returns are given by (41), then the numerator of the marginal tax rate formula (12) is*

$$1 - \sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i(E)} = 1 + \left(\frac{f_E^1(w)}{f_E(w)} - \alpha(\rho) \right) \bar{\xi} - (1 - \varepsilon_h(E)), \quad (42)$$

where $\bar{\xi}$ is given in (39).

Unsurprisingly, (42) collapses back to (39) if $\varepsilon_h = 1$ and externalities disappear. Otherwise, the optimal adjustment in (42) can be transparently decomposed into two parts: the first, local one (which varies across wage (income) levels and therefore affects the progressivity of the tax) is exactly the same as in Corollary 1. The second, new component $1 - \varepsilon_h(E)$ is of a global nature, since it uniformly scales up or down marginal keep shares $1 - T'(y)$ independent of y . In particular, if $\varepsilon_h(E) < (>)1$, we have decreasing (increasing) returns to scale and marginal tax rates are scaled up relative to an economy with constant returns to scale. This makes individuals internalize the externalities from non-constant returns to scale, which are only a function of aggregate output and therefore independent its sectoral composition.

The first correction component only depends on properties of the inner constant returns to scale production function \tilde{Y} (since this is what drives the relative returns effects), so it has exactly the same structure as in the preceding subsection. For instance, at wage levels such that $f_E^1(w)/f_E(w) = \alpha(\rho)$, there are again no relative return effects from a variation in the marginal tax rate, so the optimal adjustment equals the Pigouvian correction $1 - \varepsilon_h$, which only depends on the properties of the outer function $h(\tilde{Y})$. Finally, the adjustment in this direction never vanishes when $\varepsilon_h \neq 1$. This is because, in this setting, there exists no direction of both no relative return effects and no externalities. As before, the direction of no relative return effects, in terms of E , is $(\rho, 1)$, whereas the direction of no externalities, in terms of *incomes* is $(-1, 1)$ (since $\tau_p^1 = \tau_p^2 = 1 - \varepsilon_h$), which translates into $(-1/r_1, 1/r_2)$ in terms of E and points into a different quadrant.¹⁴ Hence, β has full rank and both consistency constraints are needed in this example.

5.3 A Pure Resource Transfer Activity

In the preceding subsection, we allowed for aggregate externalities, but the sectoral composition of incomes was still governed at the aggregate level by the constant returns to scale income shares $\alpha(\rho)$ and $1 - \alpha(\rho)$. Let us next consider the opposite case: $Y(E)$ exhibits constant returns to scale, but the aggregate income share of sector 1 may not necessarily be given by $\alpha(\rho) = Y_1(E)E_1/Y(E)$. Instead, we consider a general sectoral income decomposition $a(E)$ and $1 - a(E)$ across the two sectors. If $a(E) \neq \alpha(\rho)$, effort in one activity is underpaid relative to its marginal product, and the other is overpaid.

¹⁴Moreover, the signs of the changes ΔE_1 and ΔE_2 in E_1 and E_2 induced by a variation in the marginal tax rate at any given income are the same. Hence, $(\Delta E_1, \Delta E_2)$ can never point in the direction of zero externalities and the Pigouvian correction $1 - \varepsilon_h$ in (42) never disappears unless $\varepsilon_h = 1$.

To illustrate how this affects optimal tax policy, we consider here the extreme example where $Y(E) = E_2$, so total output only depends on effort in activity 2, but $a(E) = a(E_1)$ is an increasing function. Then effort in activity 1 is pure “stealing” of (or getting credit for) output produced in the other activity, generating income at the expense of others without adding anything to aggregate resources. In Appendix C, we show how our general formulation can be used to allow for all intermediate cases where effort in an activity may both contribute to output and increase its income share (the latter possibly decoupled from social marginal products), with similar results.

Since the social marginal product of activity 1 is zero, we have $\alpha = 0$ and activity 1 is overpaid for any $a > 0$. In particular, $\tau_p^1 = 1$ and $\tau_p^2 = -a/(1-a)$: Because activity 1 is pure rent-seeking, the Pigouvian tax is 100%, whereas activity 2 generates positive externalities (it increases the returns $r_1 = a(E_1)E_2/E_1$ to activity 1), so it commands a Pigouvian subsidy. Moreover, we have $a\tau_p^1 + (1-a)\tau_p^2 = 0$, i.e. the Pigouvian correction is zero at the aggregate level. This is because externalities here only result from the distribution of incomes across sectors, not from overall output (in contrast to the preceding subsection).

Defining the elasticity of a as $\varepsilon_1(E_1) \equiv (\partial a(E_1)/\partial E_1)E_1/a(E_1)$ yields

Proposition 6. *If $a(E) = a(E_1)$ and $Y(E) = E_2$, then*

$$\sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = (1 - (1-a-\varepsilon_1)\bar{\xi}) \frac{f_E^1(w)}{f_E(w)} + \left(-\frac{a}{1-a} + a\bar{\xi} \right) \frac{f_E^2(w)}{f_E(w)} \quad (43)$$

with

$$\bar{\xi} = \frac{I_1 + R_1 + (C + S)/(1-a)}{a(1-a)Y + (1-\varepsilon_1)(C + S)}.$$

$\bar{\xi} \geq 0$ whenever the rent-seeking activity 1 is also the high income activity, so $I_1 \geq 0$, and Pareto weights are (weakly) higher among same-wage earners on those with a high income share in the productive activity 2, so $R_1 \geq 0$ (and recall that the denominator corresponds to an eigenvalue of the matrix \mathbf{A} , so it is positive in a stable fixed point according to Lemma 7 in Appendix B).

The terms in brackets, weighted by the local income shares, collect both the Pigouvian tax rates and the adjustments for relative return effects. In particular, as seen above, there is a Pigouvian tax of 1 on activity 1 and a subsidy $-a/(1-a)$ on activity 2. The relative return effects are also intuitive. For instance, considering activity 2, the total subsidy is always less, in absolute value, than the Pigouvian subsidy. This is not surprising: A subsidy on activity 2, raising E_2 , has no effect on the returns $r_2 = 1 - a(E_1)$ to activity 2, but it increases the returns to activity 1, so it always increases the relative returns x^1 . Since this

leads to a wasteful shift of effort into activity 1, the optimum involves an *undercorrection* relative to the Pigouvian subsidy.

As for the correction on sector 1, the relative returns adjustment depends on $\varepsilon_1 \geq 1 - a$. Intuitively, an increase in E_1 has two effects: it increases x^1 because it increases $a(E_1)$, thereby reducing r_2 and increasing r_1 . But it also affects crowding in activity 1. In particular, if $a(E_1)$ does not increase much with E_1 , an increase in E_1 in fact reduces $r_1 = a(E_1)E_2/E_1$. Therefore, x^1 increases when the elasticity ε_1 is large and vice versa, depending on whether the within-sector crowding or the across-sector stealing effect dominates. For instance, when $\varepsilon_1 < 1 - a$, then a reduction in E_1 increases the relative returns to activity 1 as the crowding effect dominates, so a tax increase would lead to a flow of effort *into* the rent-seeking activity 1. The optimal correction on activity 1-intensive parts of the income distribution is therefore also less than the Pigouvian correction.

An alternative way of writing (43) is

$$\sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\zeta_i}{r_i} = \frac{1}{1-a} \left(\frac{f_E^1(w)}{f_E(w)} - a \right) - \left(\frac{f_E^1(w)}{f_E(w)} (1 - \varepsilon_1) - a \right) \bar{\zeta}, \quad (44)$$

where the first term collects the Pigouvian corrections (recall that $\tau_p^1 = 1$, which gets multiplied by the usual difference between the local and aggregate income shares of activity 1) and the second the relative return effects. This reveals that this particular example remains a two-dimensional problem in terms of the outer problem, since the directions of no externalities and no relative return effects do not coincide. The direction of no externalities is $(a, 1 - a)$ in (Y^1, Y^2) -space, so the Pigouvian term disappears when $f_E^1(w)/f_E(w) = a$. On the other hand, the direction of no relative return effects, in terms of incomes, is $(a, 1 - a - \varepsilon_1)$, so the second term disappears whenever the local income share of sector 1 is $a/(1 - \varepsilon_1)$.

It is also worth pointing out that (44) has sharp implications for the optimal progressivity of the income tax schedule. In particular, under the above conditions the Pigouvian motives for taxation here lead to a *more* progressive tax schedule than in a standard Mirrlees model (for instance, the top marginal tax rate is positive), in contrast to what we found in subsection 5.1. The relative return effects then determine whether the progressivity is even more or less pronounced than in this Pigouvian benchmark.

5.4 Externalities from One Activity

Suppose all the returns only depend on aggregate effort in one activity, i.e. $r_i(E) = r_i(E_1)$ for all $i = 1, \dots, N$. A special case of this setting was analyzed in Rothschild and Scheuer

(2013a) for $N = 2$, where a rent-seeking activity 1 imposed negative externalities on both activities, but a traditional activity 2 did not impose any externalities, so $\beta_1^j < 0$ and $\beta_2^j = 0$ for $j = 1, 2$. We can use the tools developed here to solve this model for general N and any form of externalities generated by sector 1.

In particular, since $\beta_i^j = 0$ for all $i = 2, \dots, N$ and all j , the matrix β has rank one (all rows except for the first are zero) and the optimality condition (32) implies $\xi_i = 0$ for all $i \neq 1$. Intuitively, there is a whole $N - 1$ -dimensional subspace of directions of no externalities and no relative return effects, spanned by all the vectors in E -space with a zero first element. Only E_1 generates externalities and relative return effects, so the dimensionality of the outer problem reduces to one. This leads to the following result:

Proposition 7. *If $r_i(E) = r_i(E_1)$ for all $i = 1, \dots, N$, then the numerator of the marginal tax rate formula in (12) is $1 - \bar{\xi} f_E^1(w) / f_E(w)$ with*

$$\bar{\xi} = \frac{\tau_p^1 + \sum_{j=1}^{N-1} \Delta\beta_1^j (I_j + R_j) / r_1}{1 - \sum_{j=1}^{N-1} \Delta\beta_1^j (C_{j1} + S_{j1})}. \quad (45)$$

The optimal adjustment to the marginal income tax formula is $\bar{\xi}$ weighted by the local income share of the externality generating activity 1 at w . $\bar{\xi}$ in turn deviates from the Pigouvian correction τ_p^1 only if there are relative return effects, so $\Delta\beta_1^j \neq 0$ for some j . These enter in an intuitive way. For instance, suppose activity 1 generates negative externalities, so $\tau_p^1 > 0$. Then the denominator in (45) increases $\bar{\xi}$ compared to this Pigouvian correction if an increase in E_1 on average raises the relative returns to activities j with $C_{j1}, S_{j1} > 0$, and vice versa. This is because an increase in the relative returns x^j to these activities leads to a flow of effort into activity 1, since $\text{Cov}(q^1, q^j) > 0$ and $Q_j^1 > 0$ in this case. A tax on sector 1 income, through reducing E_1 and thus inducing the opposite flow of effort out of activity 1, is therefore even more desirable than based on the purely Pigouvian motives.¹⁵

The second term in the numerator of (45) further increases $\bar{\xi}$ compared to τ_p^1 if the activities whose relative returns increase in response to an increase in E_1 are also high income, low Pareto weight activities on average (so $\Delta\beta_1^j$ is positively correlated with I_j, R_j).¹⁶ Then an increase in the marginal income tax at wage levels where activity 1 is prevalent raises the returns to lower wage, high redistributive preference activities by

¹⁵Note that the denominator of (45) is always positive when the optimum involves a stable fixed point for E_1 in terms of Lemma 7 in Appendix B, since it is the eigenvalue of the matrix A associated with the unit eigenvector $(1, 0, 0, \dots, 0)'$.

¹⁶Recall that $\sum_j I_j = \sum_j R_j = 0$, so $\sum_j \Delta\beta_1^j (I_j + R_j)$ can be interpreted as N times $\text{Cov}(\Delta\beta_1^j, I_j + R_j)$ across activities j .

decreasing E_1 , thus achieving indirect redistribution. Of course, analogous results can be obtained from (45) when the tax leads to the opposite sectoral shifts, giving rise to an undercorrection at the optimum, or when activity 1 imposes positive or mixed externalities.

The special case considered in Rothschild and Scheuer (2013a) for $N = 2$ immediately emerges as

$$\bar{\xi} = \frac{\tau_p^1 + \Delta\beta_1^1(I_1 + R_1)/r_1}{1 - \Delta\beta_1^1(C + S)/r_1}$$

with C and S given by (33) and (34). If $l(w)$ is increasing, so that $C > 0$, and if we also have $I_1, R_1 > 0$ (because the externality-causing activity $i = 1$ is also a high wage and low redistributive preference activity), then an undercorrection with $\bar{\xi} < \tau_p^1$ is optimal if $\Delta\beta_1^1 < 0$ (so a decrease in E_1 increases the relative returns x^1 and leads to a perverse sectoral shift into activity 1) and an overcorrection with $\bar{\xi} > \tau_p^1$ otherwise. Note again that these results do not depend on the form of the externalities generated by activity 1; in particular, they are not confined to the special case in Rothschild and Scheuer (2013a) with only negative externalities.

5.5 Externalities Targeted at One Activity

Let us finally turn to the opposite case, where the returns to only one activity depend on E , so that $r_1(E)$ is general but $r_i(E) = r_i$ are constants for all $i = 2, \dots, N$. A simple example with $N = 2$ would be another specification of a pure resource transfer activity with $Y(E) = Y(E_1)$ and $Y^1(E) = Y(E_1) - E_2$ and $Y^2(E_2) = E_2$. Hence, all output is produced through activity 1, and activity 2 takes away some of this output one-for-one.¹⁷

Generally, $\beta_i^j = 0$ for all $j \neq 1$ and β again has rank one in this case, this time with all *columns* being zero except for the first, which has elements β_i^1 (and $\Delta\beta = \beta$). Intuitively, any movement in E -space that changes $r_1(E)$ generates both an externality and a relative return change. Conversely, since $t_p^i = -\beta_i^1 Y^1$ in this example, all the $N - 1$ dimensions of \mathbb{R}^N orthogonal to the vector $(\beta_1^1, \beta_2^1, \dots, \beta_N^1)$ are directions of both no externalities and no relative return effects because changes of E in these directions leave $r_1(E)$ unchanged. As a result, there is only one effective consistency constraint in the outer problem here, which is a β_i^1 -weighted sum of the original N constraints (8).

Using this, (32) immediately implies $\xi_i/\beta_i^1 = \xi_1/\beta_1^1$ for all i and therefore the following result:

¹⁷For $N = 2$, this is a special case of the example considered in Appendix C with $a(E) = 1 - E_2/Y(E_1)$. But we could imagine a more general model with $N - 1$ stealing sectors $i = 2, \dots, N$ and $Y^1(E) = Y(E_1) - \sum_{i=2}^N (1 + k_i)E_i$, $Y^i(E_i) = E_i$ for $i \geq 2$, where these activities differ in terms of the share k_i of resources lost in the process of transferring them.

Proposition 8. *If $r_i(E)$ is fixed for all $i \neq 1$, then the optimal adjustment term in (12) is*

$$\sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \frac{\zeta_i}{r_i} = \sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \frac{\beta_i^1}{r_i} \bar{\zeta} \quad \text{with} \quad \bar{\zeta} = \frac{-Y^1 + I_1 + R_1}{1 - \sum_{i=1}^N \beta_i^1 (C_{1i} + S_{1i})}. \quad (46)$$

Since both the externalities and the relative return effects induced by a change in E_i are scaled by the magnitude of β_i^1 , the optimal correction (in terms of income) in each dimension i is proportional to β_i^1/r_i . Hence, the adjustment factor vanishes whenever the vector of local income shares at w is orthogonal to the vector of these magnitudes, i.e. $\sum_i (f_E^i(w)/f_E(w))(\beta_i^1/r_i) = 0$. Intuitively, this is a wage level at which a variation in the marginal income tax rate leads to changes in E that leave $r_1(E)$ unaffected, so the optimal marginal tax rate is as if all returns were fixed locally.

Otherwise, suppose, for instance, that all activities increase $r_1(E)$, so $\beta_1^1(E) > 0$ and $\sum_i (f_E^i(w)/f_E(w))(\beta_i^1/r_i) > 0$ for all w . The first term in the numerator of $\bar{\zeta}$ in (46) captures the Pigouvian subsidy for these positive externalities, since

$$-\sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \frac{\beta_i^1}{r_i} Y^1 = \sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \tau_p^i.$$

In this case, the Pigouvian correction alone would lead to a negative adjustment term and therefore lower marginal tax rates in (12). The denominator and the second term $I_1 + R_1$ in the numerator capture the deviation from this Pigouvian adjustment due to the relative return effects from the increase in r_1 induced by the subsidy.¹⁸

The term $I_1 + R_1$ in the numerator of $\bar{\zeta}$ captures the direct effect of the rise in r_1 : If activity 1 is in fact a relatively high wage activity with little redistributive weight, so $I_1 + R_1 > 0$, then the increase in r_1 is not desirable for distributional reasons, which is why the optimal subsidy is *less* than the Pigouvian subsidy, and vice versa. The denominator is a multiplier term coming from the indirect effects of the increase in r_1 through the induced sectoral reallocations of effort. It is easiest to understand when $N = 2$, in which case it reduces to $1 - (\beta_1^1/r_1 - \beta_2^1/r_2)(C + S)$ with C and S given in (33) and (34). Intuitively, the increase in r_1 , raising the relative returns to activity 1, will always lead to a flow of effort from activity 2 to 1. Whether this flow reinforces or mitigates the original increase in r_1 depends on the relative magnitudes of the externalities from E_1 and E_2 . In particular, since the flow increases E_1 and reduces E_2 , there will be a further increase in r_1 if β_1^1 is large compared to β_2^1 , and a reduction otherwise. The denominator scales the direct effects in

¹⁸The denominator of $\bar{\zeta}$ in (46) is positive if the optimum involves a stable fixed point of E according to Lemma 7 in Appendix B, since it is an eigenvalue of the matrix A associated with the eigenvector $(\beta_1^1, \beta_2^1, \dots, \beta_N^1)'$.

the numerator to account for these indirect multiplier effects of the sectoral shifts on r_1 .

6 Conclusion

As suggested by the examples in the preceding section, the framework developed here is flexible enough to handle a wide variety of applications. It is important to emphasize, however, that these examples are not exhaustive: the optimal tax formula (12) and the characterization of the correction term in that formula through condition (32) are fully general and could be used to explore other special cases in future research.

While adapting these formulas for applied policy work will be non-trivial and is beyond the scope of this paper, they provide useful insights into the nature of evidence that would be required to implement them. For instance, in the pure resource transfer example discussed in Section 5.3, the Pigouvian component of the correction would be entirely pinned down by the aggregate income share accruing to the transfer activity. The divergence between this and the optimal correction in turn only depends on the elasticity of this income share with respect to aggregate effort in this activity: If this elasticity is low, the within-sector crowding effects dominate and the transfer activity itself bears the bulk of the externalities. As such, the Pigouvian tax induces a perverse shift of effort into this activity, and the optimal correction falls short of it. If the elasticity is high, the externalities are borne primarily by the productive activity, so a tax induces the opposite, beneficial shift and an *overcorrection* is optimal. Thus, information on these income shares and elasticities would be of direct use for optimal policy design.

More generally, the key applied-policy lesson is that in settings with externalities, the simple Pigouvian wedge between the private and social marginal returns of the average worker at a given income is not sufficient for determining the optimal corrective adjustment to the marginal tax rate at that income level. Policy makers need to know not just *who* is over- or underpaid and by how much, but also *on whom* the resulting externalities are imposed.

References

- ATKINSON, A., T. PIKETTY, AND E. SAEZ (2011): "Top Incomes in the Long-Run of History," *Journal of Economic Literature*, 49, 3–71.
- BESLEY, T., AND M. GHATAK (2013): "Bailouts and the Optimal Taxation of Bonus Pay," *American Economic Review, Papers and Proceedings*, 103, 163–167.

- BIAIS, B., T. FOUCAULT, AND S. MOINAS (2011): "Equilibrium High Frequency Trading," *Mimeo, TSE*.
- BOOTHBY, W. (1986): *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, Orlando, Florida, 2 edn.
- CHONÉ, P., AND G. LAROQUE (2010): "Negative Marginal Tax Rates and Heterogeneity," *American Economic Review*, 100, 2532–2547.
- CREMER, H., F. GAHVARI, AND N. LADOUX (1998): "Externalities and Optimal Income Taxation," *Journal of Public Economics*, 70, 343–364.
- DIAMOND, P. (1973): "Consumption Externalities and Imperfect Corrective Pricing," *Bell Journal of Economics and Management Science*, 4, 526–538.
- DIAMOND, P. (1998): "Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Tax Rates," *American Economic Review*, 88, 83–95.
- DIXIT, A. (1985): "Tax Policy in Open Economies," in *Handbook of Public Economics*, ed. by A. Auerbach, and M. Feldstein, pp. 2475–2490. North Holland, Amsterdam.
- FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. MIT Press, Cambridge.
- GLODE, V., AND R. LOWERY (2012): "Compensating Financial Experts," *Mimeo, University of Pennsylvania*.
- KLEVEN, H. J., C. T. KREINER, AND E. SAEZ (2009): "The Optimal Income Taxation of Couples," *Econometrica*, 77, 537–560.
- KOPCZUK, W. (2003): "A Note on Optimal Taxation in the Presence of Externalities," *Economics Letters*, 80, 81–86.
- LOCKWOOD, B., C. NATHANSON, AND G. WEYL (2012): "Taxation and the Allocation of Talent," *Mimeo, Harvard University and University of Chicago*.
- MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.
- PIKETTY, T., E. SAEZ, AND S. STANTCHEVA (2013): "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, forthcoming.

- 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*.
- (2013a): “Optimal Taxation with Rent-Seeking,” *Mimeo, Stanford University and Wellesley College*.
- (2013b): “Redistributive Taxation in the Roy Model,” *Quarterly Journal of Economics*, 128, 623–668.
- ROY, A. (1951): “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 3, 135–146.
- SADKA, E. (1978): “On the Optimal Taxation of Consumption Externalities,” *Quarterly Journal of Economics*, 92.
- SAEZ, E. (2001): “Using Elasticities to Derive Optimal Tax Rates,” *Review of Economic Studies*, 68, 205–229.
- SANDMO, A. (1975): “Optimal taxation in the presence of externalities,” *The Swedish Journal of Economics*, 77, 86–98.
- SCHEUER, F. (2013): “Entrepreneurial Taxation with Endogenous Entry,” *forthcoming, American Economic Journal: Economic Policy*.
- 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*.

A Proofs

A.1 Proof of Lemma 1

Using (2) for $\theta = \theta^l$ and homogeneity of degree one of m , we have

$$q(\theta) \in \arg \min_{p \in \Delta^{N-1}} m \left(\frac{p_1 y(\theta)}{\theta_1 r_1(E)}, \dots, \frac{p_N y(\theta)}{\theta_N r_N(E)} \right) = \arg \min_{p \in \Delta^{N-1}} m \left(\frac{p_1}{\theta_1 r_1(E)}, \dots, \frac{p_N}{\theta_N r_N(E)} \right).$$

The result in (3) follows from $w(\theta) \equiv y(\theta)/l(\theta)$ and $l(\theta) \equiv m(e(\theta)) = y(\theta) m \left(\frac{q_1(\theta)}{\theta_1 r_1(E)}, \dots, \frac{q_N(\theta)}{\theta_N r_N(E)} \right)$.

A.2 Proof of Lemma 2

The proof is analogous to that of Lemma 1 in Rothschild and Scheuer (2013b) for $N = 2$.

A.3 Proof of Proposition 1

Putting multipliers λ on (9), $\xi_i \lambda$ on the consistency constraints (8), and $\hat{\eta}(w)\lambda$ on (7), the Lagrangian corresponding to (6)-(9) is, after integrating by parts (7),

$$\begin{aligned} \mathcal{L} = & \int_{\underline{w}_E}^{\bar{w}_E} V(w)\psi_E(w)dw - \int_{\underline{w}_E}^{\bar{w}_E} V(w)\hat{\eta}'(w)\lambda dw + \int_{\underline{w}_E}^{\bar{w}_E} u_l(c(V(w),l(w)),l(w))\frac{l(w)}{w}\hat{\eta}(w)\lambda dw \\ & + \sum_{i=1}^N \xi_i \lambda \left[E_i - \frac{1}{r_i(E)} \int_{\underline{w}_E}^{\bar{w}_E} wl(w)f_E^i(w)dw \right] + \lambda \int_{\underline{w}_E}^{\bar{w}_E} (wl(w) - c(V(w),l(w)))f_E(w)dw. \end{aligned} \quad (47)$$

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

$$\hat{\eta}'(w)\lambda = \psi_E(w) - \lambda f_E(w)\frac{1}{u_c(w)} + \hat{\eta}(w)\lambda \frac{u_{cl}(w)}{u_c(w)} \frac{l(w)}{w}. \quad (48)$$

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

$$\eta'(w) = \psi_E(w)\frac{u_c(w)}{\lambda} - f_E(w) + \eta(w)\frac{u_{cc}(w)c'(w) + u_{cl}(w)l'(w) + u_{cl}(w)l(w)/w}{u_c(w)}. \quad (49)$$

Using the first order condition corresponding to the incentive constraint (11),

$$u_c(w)c'(w) + u_l(w)l'(w) + u_l(w)\frac{l(w)}{w} = 0, \quad (50)$$

the fraction in (49) can be written as $-(\partial MRS(w)/\partial c)y'(w)/w$, where $M(c,l) \equiv -u_l(c,l)/u_c(c,l)$ is the marginal rate of substitution between effort and consumption and $MRS(w) \equiv M(c(w),l(w))$, so (with a slight abuse of notation) $\partial MRS(w)/\partial c$ stands short for $\partial M(c(w),l(w))/\partial c$. Substituting in (49) and rearranging yields

$$-\frac{\partial MRS(w)}{\partial c}l(w)\frac{y'(w)}{y(w)}\eta(w) = f_E(w) - \psi_E(w)\frac{u_c(w)}{\lambda} + \eta'(w). \quad (51)$$

Integrating this ODE gives

$$\begin{aligned} \eta(w) &= \int_w^{\bar{w}_E} \left(f_E(s) - \psi_E(s)\frac{u_c(s)}{\lambda} \right) \exp \left(\int_w^s \frac{\partial MRS(t)}{\partial c} l(t) \frac{y'(t)}{y(t)} dt \right) ds \\ &= \int_w^{\bar{w}_E} \left(1 - \frac{\psi_E(s)}{f_E(s)} \frac{u_c(s)}{\lambda} \right) \exp \left(\int_w^s \left(1 - \frac{\varepsilon^u(t)}{\varepsilon^c(t)} \right) \frac{dy(t)}{y(t)} \right) f_E(s) ds, \end{aligned} \quad (52)$$

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

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

$$\lambda w f_E(w) \left(1 - \frac{MRS(w)}{w} \right) - \lambda w \sum_{i=1}^N \frac{\xi_i}{r_i(E)} f_E^i(w) = -\hat{\eta}(w)\lambda \left[\frac{(-u_{cl}(w)u_l(w)/u_c(w) + u_{ll}(w))l(w)}{w} + \frac{u_l(w)}{w} \right],$$

which after some algebra can be rewritten as

$$wf_E(w) \left(1 - \frac{MRS(w)}{w}\right) - w \sum_{i=1}^N \frac{\xi_i}{r_i(E)} f_E^i(w) = \eta(w) \left(\frac{\partial MRS(w)}{\partial l} \frac{l}{w} + \frac{MRS(w)}{w} \right), \quad (53)$$

where $\partial MRS(w)/\partial l$ again stands short for $\partial M(c(w), l(w))/\partial l$. With $MRS(w)/w = 1 - T'(y(w))$ from the first order condition of the workers, this becomes

$$1 - \sum_{i=1}^N \frac{\xi_i}{r_i(E)} \frac{f_E^i(w)}{f_E(w)} = (1 - T'(y(w))) \left[1 + \frac{\eta(w)}{wf_E(w)} \left(1 + \frac{\partial MRS(w)}{\partial l} \frac{l}{MRS(w)} \right) \right]. \quad (54)$$

Simple algebra again shows that $1 + \partial \log MRS(w)/\partial \log l = (1 + \varepsilon^u(w))/\varepsilon^c(w)$, so that the result follows from (52) and (54).

A.4 Proof of Lemma 3

(i) For C_{kj} , this follows from

$$\sum_{j=1}^N r_j(E) C_{kj}(E) = \int_{\underline{w}_E}^{\bar{w}_E} w^2 l'(w) \sum_{j=1}^N \text{Cov} \left(q_E^j, q_E^k \mid w \right) dw = 0$$

because $\sum_{j=1}^N \text{Cov} \left(q_E^j, q_E^k \mid w \right) = \text{Cov} \left(\sum_{j=1}^N q_E^j, q_E^k \mid w \right) = \text{Cov} \left(1, q_E^k \mid w \right) = 0$ for all w . For S_{kj} , we prove the result by showing that $\sum_{j=1}^N Q_k^j(x_E(\phi)) = 0$ for all $\phi \in \Phi$. For this, use (28) and (29) to write

$$0 = \frac{\partial 1}{\partial(\theta_k r_k(E))} = \sum_{j=1}^N \frac{\partial q_E^j(\phi)}{\partial(\theta_k r_k(E))} = \sum_{j=1}^N \frac{\partial Z_j(x_E(\phi))}{\partial(\theta_k r_k(E))} \Omega_j(\zeta(x_E(\phi))) + \frac{1}{\theta_N r_N(E)} \sum_{j=1}^N Q_k^j(x_E(\phi)) \quad \forall \phi.$$

Hence, showing that $\sum_j \Omega_j \partial Z_j / \partial(\theta_k r_k(E)) = 0$ will complete the proof. Using (28), we have

$$Z_j(x_E(\phi)) \equiv \frac{r_j(E)\theta_j}{w} = r_j(E)\theta_j \min_{p \in \Delta^{N-1}} m \left(\frac{p_1}{r_1(E)\theta_1}, \dots, \frac{p_N}{r_N(E)\theta_N} \right), \quad (55)$$

so

$$\frac{\partial Z_j}{\partial(r_k \theta_k)} = \frac{\delta_{kj}}{w} - r_j \theta_j m_k(e/y) \frac{q^k}{(r_k \theta_k)^2}, \quad (56)$$

where m_k denotes the (homogeneous of degree zero) partial derivative of m w.r.t. its k -th argument. Note that the first order conditions for the minimization in (55) are $m_k(e/y)/(r_k \theta_k) = m_N(e/y)/(r_N \theta_N)$ for all $k = 1, \dots, N$, which implies

$$\frac{1}{w} = \frac{m(e)}{y} = m(e/y) = \sum_{k=1}^N m_k(e/y) \frac{q^k}{r_k \theta_k} = \frac{m_N(e/y)}{r_N \theta_N} \sum_{k=1}^N q^k = \frac{m_N(e/y)}{r_N \theta_N} = \frac{m_k(e/y)}{r_k \theta_k} \quad \forall k, \quad (57)$$

where the second equality uses linear homogeneity of m , the third uses Euler's theorem, and the fourth and sixth the first order conditions. Substituting this in (56) and using $\Omega_j = e_j/m(e) = e_j w/y$ yields

$$\begin{aligned} \sum_j \Omega_j \frac{\partial Z_j}{\partial(\theta_k r_k)} &= \sum_j \frac{e_j w}{y} \left(\frac{\delta_{jk}}{w} - \frac{r_j \theta_j}{w} \frac{q^k}{r_k \theta_k} \right) = \sum_j \frac{e_j}{y} \left(\delta_{jk} \frac{r_j \theta_j}{r_j \theta_j} - q^k \frac{r_j \theta_j}{r_k \theta_k} \right) = \sum_j \frac{y_j}{y} \left(\frac{\delta_{jk}}{r_k \theta_k} - \frac{q^k}{r_k \theta_k} \right) \\ &= \sum_j \frac{1}{r_k \theta_k} (q^j \delta_{jk} - q^j q^k) = \frac{1}{r_k \theta_k} (q^k - q^k) = 0, \end{aligned}$$

where the third equality uses $\delta_{jk}/r_j \theta_j = \delta_{jk}/r_k \theta_k$. The remaining steps are algebra and establish the result.

(ii) $\sum_{k=1}^N S_{kj}(E) = 0$ follows from (31) and $\sum_{k=1}^N Q_k^j(x_E(\phi)) x_E^k(\phi) = 0$ for all ϕ . To see the latter, use (29) and

$$\sum_{k=1}^N \frac{\partial \zeta_l(x_E(\phi))}{\partial(r_k(E)\theta_k)} x_E^k(\phi) r_N(E) \theta_N = \sum_{k=1}^N \frac{\partial \zeta_l(x_E(\phi))}{\partial(r_k(E)\theta_k)} r_k(E) \theta_k = 0 \quad \forall l$$

by the zero-homogeneity of ζ and Euler's theorem. $\sum_{k=1}^N C_{kj}(E) = 0$ follows from (26) and an analogous argument to part (i).

A.5 Proof of Proposition 2

From (23), $\bar{n}' \bar{t}_p = \bar{n}' \beta \bar{Y}$, where \bar{Y} denotes the column vector of aggregate sectoral incomes $Y^i(E)$. By definition, $\Delta \beta = \beta (\mathbf{I}_N - \mathbf{O}_N)$, where \mathbf{O}_N is matrix with $(i, j)^{\text{th}}$ element δ_{Nj} (i.e., with ones in the last row and zeros otherwise). The "if" is thus immediate. For "only if", observe that the last column of $\mathbf{I}_N - \mathbf{O}_N$ is zero and let \mathbf{D} denote the matrix whose first $N - 1$ columns coincide with $\mathbf{I}_N - \mathbf{O}_N$ and whose N^{th} column is \bar{Y} . Then $\bar{n}' \Delta \beta = 0$ and $\bar{n}' \bar{t}_p = 0$ only if $\bar{n}' \beta \mathbf{D} = 0$. Since $\bar{Y} \geq 0$ with at least one strictly positive entry, \mathbf{D} is non-singular. Hence, $\bar{n}' \beta \mathbf{D} = 0$ only if $\bar{n}' \beta = 0$.

A.6 Proof of Proposition 3

Since $w_E(\theta)$ and $q_E(\theta)$ depend on E only through the returns vector $r(E)$, this vector is a sufficient statistic for individual decisions given any $l(w)$ and $V(w)$, and hence for the solution to the inner problem. β has the same rank, $N - K$, as the matrix of partial derivatives $Dr(\cdot)$, as $\ln(\cdot)$ is a diffeomorphism. By the Constant Rank Theorem (Boothby, 1986, Theorem 7.1), there exist open neighborhoods $U_E \subset \mathbb{R}^N$ of E^* and $U_r \subset \mathbb{R}^N$ of $r(E^*)$ and diffeomorphisms G from U_E onto an open subset of \mathbb{R}^N and H from U_r onto an open subset of \mathbb{R}^N such that $H(r(G^{-1}(x_1, \dots, x_N))) = (x_1, \dots, x_{N-K}, 0, \dots, 0)$. Defining $\rho \equiv (x_1, \dots, x_{N-K})$, we have $r(G^{-1}(\rho, x_{N-K+1}, \dots, x_N)) = H^{-1}(\rho, 0, \dots, 0)$, so ρ is sufficient for r .

To find the consistency constraints associated with ρ , let $\mathcal{E}(r(E); l(\cdot))$ denote the vector of right-hand sides of (8). Then the i^{th} consistency constraint, $i = 1, \dots, N - K$ is $\rho_i = G_i(\mathcal{E}(H^{-1}(\rho, 0, \dots, 0); V(\cdot), l(\cdot)))$, i.e., the i^{th} component of $G(E) = G(\mathcal{E}(r(E); l(\cdot)))$, written in terms of ρ .

A.7 Proof of Lemma 5

Dropping the arguments E , the optimality conditions (32) can be written for $N = 2$ as

$$\mathbf{A} \bar{\zeta} = \bar{t}_p + \begin{pmatrix} \Delta \beta_1^1 \\ \Delta \beta_2^1 \end{pmatrix} (\mathbf{I}_1 + \mathbf{R}_1). \quad (58)$$

Since $(\Delta\beta_1^1, \Delta\beta_2^1)'$ is an eigenvector of \mathbf{A} , it is also an eigenvector of \mathbf{A}^{-1} (with associated eigenvalue $1/\gamma_2$), and we can write (58) as

$$\vec{\zeta} = \mathbf{A}^{-1}\vec{t}_p + \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} \frac{I_1 + R_1}{\gamma_2}.$$

Moreover, defining the eigenbasis

$$\mathbf{B} \equiv \begin{pmatrix} r_1 & \Delta\beta_1^1 \\ r_2 & \Delta\beta_2^1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} a \\ b \end{pmatrix} \equiv \mathbf{B}^{-1}\vec{t}_p,$$

we can write $\vec{t}_p = a \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} + b \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix}$. Using this and $\frac{1}{\gamma_2} = 1 - \frac{1}{\gamma_2} \left(\frac{\Delta\beta_2^1}{r_2} - \frac{\Delta\beta_1^1}{r_1} \right) (C + S)$, we have

$$\begin{aligned} \mathbf{A}^{-1}\vec{t}_p &= \mathbf{A}^{-1}\mathbf{B} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mathbf{A}^{-1} \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} & \mathbf{A}^{-1} \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = a \begin{pmatrix} r_1 \\ r_2 \end{pmatrix} + \frac{b}{\gamma_2} \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} \\ &= \vec{t}_p - b \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} + \frac{b}{\gamma_2} \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} = \vec{t}_p - \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} \frac{b}{\gamma_2} \left(\frac{\Delta\beta_2^1}{r_2} - \frac{\Delta\beta_1^1}{r_1} \right) (C + S). \end{aligned}$$

Hence,

$$\vec{\zeta} = \vec{t}_p + \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} \frac{I_1 + R_1 - b(\Delta\beta_2^1/r_2 - \Delta\beta_1^1/r_1)(C + S)}{\gamma_2}. \quad (59)$$

Finally, note that the second row of \mathbf{B}^{-1} is $(-1/r_1, 1/r_2) / (\Delta\beta_2^1/r_2 - \Delta\beta_1^1/r_1)$, so

$$b = - \left(\frac{t_p^1}{r_1} - \frac{t_p^2}{r_2} \right) / \left(\frac{\Delta\beta_2^1}{r_2} - \frac{\Delta\beta_1^1}{r_1} \right).$$

Substituting in (59) yields the result.

A.8 Proof of Lemma 6

Since $r_i(E) = Y_i(E)$ when there are no externalities, $Y(E) = \sum_{i=1}^N r_i(E)E_i = \sum_{i=1}^N Y_i(E)E_i$. The lemma then follows from Euler's Theorem.

A.9 Proof of Proposition 4

Observe first that $\beta_i^3 = 0$ for all i , so the third row of $\Delta\beta$ is zero. Using this together with $\vec{t}_p = 0$ in (32) immediately implies $\zeta_3 = 0$. Straightforward calculations yield $\beta_1^1(E) = -\beta_2^1(E)/\rho = Y_1'(\rho)/(E_2Y_1(\rho))$, $\beta_2^1(E) = -\beta_2^2(E)/\rho = Y_2'(\rho)/(E_2Y_2(\rho))$, $\Delta\beta_1^1(E) = -1/(E_1\hat{\sigma}(\rho))$, and $\Delta\beta_2^1(E) = 1/(E_2\hat{\sigma}(\rho))$, where $-1/\hat{\sigma}(\rho) = \rho(Y_1'(\rho)/Y_1(\rho) - Y_2'(\rho)/Y_2(\rho))$ is the substitution elasticity of \hat{Y} . Hence, the second row of $\Delta\beta$ is $-\rho$ times the first row, which implies $\zeta_2 = -\rho\zeta_1$. We can therefore use the first row of the system (32) to solve for ζ_1 , which (using $\beta_1^3 = 0$) yields $\zeta_1/r_1 = -(1 - \hat{\alpha})\hat{\zeta}$ where $\hat{\zeta}$ is given in (38), and $\zeta_2/r_2 = -\rho\zeta_1/r_1 = \hat{\alpha}\hat{\zeta}$. Finally, substituting these two and $\zeta_3 = 0$ in the adjustment term delivers (37).

A.10 Proof of Proposition 5

Tedious algebra yields

$$\beta_1^1(E)E_1 = -\frac{1-\alpha(\rho)}{\sigma(\rho)} - \alpha(\rho)(1-\varepsilon_h(E)), \quad \beta_2^1(E)E_2 = \frac{1-\alpha(\rho)}{\sigma(\rho)} - (1-\alpha(\rho))(1-\varepsilon_h(E)),$$

$$\beta_1^2(E)E_1 = \frac{\alpha(\rho)}{\sigma(\rho)} - \alpha(\rho)(1-\varepsilon_h(E)), \quad \beta_2^2(E)E_2 = -\frac{\alpha(\rho)}{\sigma(\rho)} - (1-\alpha(\rho))(1-\varepsilon_h(E)),$$

so $\Delta\beta_1^1(\rho)E_1 = -\Delta\beta_2^2(\rho)E_2 = -1/\sigma(\rho)$. Moreover, $\tau_p^1(E) = \tau_p^2(E) = 1 - \varepsilon_h(E)$. Substituting in (36) yields $\xi_1/r_1 = 1 - \varepsilon_h(E) - (1 - \alpha(\rho))\bar{\xi}$ and $\xi_2/r_2 = 1 - \varepsilon_h(E) + \alpha(\rho)\bar{\xi}$, and using this in (12) yields (42).

A.11 Proof of Proposition 6

Proposition 6 is a direct Corollary of Proposition 9 in Appendix C for $\varepsilon_2 = \alpha = 0$.

A.12 Proof of Proposition 7

We can use the first row of (32) to explicitly solve for ξ_1 , using $\xi_i = 0$ for all $i \neq 1$:

$$\xi_1 = \frac{t_p^1 + \sum_{j=1}^{N-1} \Delta\beta_1^j(I_j + R_j)}{1 - \sum_{j=1}^{N-1} \Delta\beta_1^j(C_{j1} + S_{j1})}$$

where $t_p^1 = -\sum_{j=1}^N \beta_1^j Y^j$ and C_{j1} and S_{j1} are given in (26) and (31). This immediately yields the result.

A.13 Proof of Proposition 8

Using $\xi_i = \zeta_1 \beta_i^1 / \beta_1^1$, we can use the first equation in the system (32) to solve for ζ_1 :

$$\zeta_1 = \left(t_p^1 + \beta_1^1(I_1 + R_1) \right) / \left(1 - \sum_{i=1}^N \beta_i^1 (C_{1i} + S_{1i}) \right)$$

Again using $\xi_i = \zeta_1 \beta_i^1 / \beta_1^1$ and $t_p^i = t_p^1 \beta_i^1 / \beta_1^1$ delivers the result.

B Eigenvalues and Stability

For any given vector E , and holding for instance the schedule $l(w)$ fixed, the right-hand sides of the system of consistency constraints (8) yield some implied vector of sectoral efforts $\mathcal{E}(E)$, and the constraints require that the optimal E is a fixed point of this mapping: $E = \mathcal{E}(E)$. It is reasonable to assume that the optimal E is in fact a stable fixed point of this mapping, since otherwise we would have no reason to expect that it will be reached when the government offers the optimal tax schedule $T(y)$.

However, as discussed in detail in section 3.3, the variation underlying (32) is not keeping the schedule $l(w)$ fixed, so the appropriate notion of stability of the fixed point needs to account for the schedule variation in subshift 2 when E changes. Formally, suppose we start from some optimal vector E^*

and schedule $l^*(w)$ and move locally away from E^* to E . The resulting average effort change at w is $\sum_{i=1}^N l^{*'}(w)w\delta_{E^*}^i(w)(E_i - E_i^*)$ by (16) and (17). Our variation constructs an effort schedule $l_E(w)$ in subshift 2 by making the negative of this adjustment to $l^*(w)$ at each w and for each E , i.e.

$$l_E(w) \equiv l^*(w) - \sum_{i=1}^N l^{*'}(w)w\delta_{E^*}^i(w)(E_i - E_i^*). \quad (60)$$

Note that the adjustments to the effort schedule underlying $l_E(w)$ are linear in E . If there is no bunching at the original optimum, so $y^*(w)$ is increasing, we know that $y_E(w) \equiv wl_E(w)$ will be increasing in w as well for E close to E^* , so it will be implementable with some nonlinear income tax schedule $T_E(y)$. This tax schedule is such that average effort at each w is unchanged when varying E close to E^* .

We are interested in the stability of the optimal fixed point E^* when the government offers this E -contingent tax schedule. Let $\mathcal{E}_i(E) = \int_{\underline{w}_E}^{\bar{w}_E} y_E(w)f_E^i(w)dw/r_i(E)$ and imagine a dynamic system with

$$\dot{E}_i = \mathcal{E}_i(E) - E_i, \quad i = 1, \dots, N. \quad (61)$$

Denoting the Jacobian of the right-hand side of this system, $\mathcal{E}(E) - E$, by \mathbf{J} , stability of the fixed point E^* requires the real parts of all eigenvalues of \mathbf{J} to be negative. Since $\mathcal{E}(E) - E$ is the negative of the consistency constraints (8), $\mathbf{J} = -\mathbf{A}$ by our derivation of the consistency constraint effects in section 3.3. This leads to the following result:

Lemma 7. *A fixed point E of the system (61) is stable if and only if all eigenvalues of the matrix $\mathbf{A} = \mathbf{I}_N - \Delta\beta(\mathbf{C} + \mathbf{S})$ in (32) have positive real parts.*

C General Sectoral Income Shares

C.1 The General Case

Let $Y(E)$ have constant returns to scale and $Y^1(E) = a(E)Y(E)$ and $Y^2(E) = (1 - a(E))Y(E)$, so that, by linear returns,

$$r_1(E) = a(E)Y(E)/E_1 \quad \text{and} \quad r_2(E) = (1 - a(E))Y(E)/E_2. \quad (62)$$

Defining

$$\varepsilon_1(E) \equiv \frac{\partial a(E)}{\partial E_1} \frac{E_1}{a(E)} \quad \text{and} \quad \varepsilon_2(E) \equiv \frac{\partial(1 - a(E))}{\partial E_2} \frac{E_2}{1 - a(E)}$$

yields:

Proposition 9. *If $N = 2$, $Y(E)$ has constant returns to scale and private returns are given by (62), then the adjustment to the marginal tax rate formula (12) is*

$$\sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = \frac{f_E^1(w)}{f_E(w)} \frac{a - \alpha}{a} + \frac{f_E^2(w)}{f_E(w)} \frac{\alpha - a}{1 - a} + \left(\frac{f_E^1(w)}{f_E(w)} (\varepsilon_1 - (1 - a)) + \frac{f_E^2(w)}{f_E(w)} (a - \varepsilon_2) \right) \bar{\xi}$$

with

$$\bar{\xi} = \frac{I_1 + R_1 + \frac{a - \alpha}{a(1 - a)}(C + S)}{a(1 - a)Y + (1 - \varepsilon_1 - \varepsilon_2)(C + S)}. \quad (63)$$

Proof. We have

$$\beta_1^1(E)E_1 = -(1 - \alpha(\rho)) + \varepsilon_1(E), \quad \beta_2^1(E)E_2 = 1 - \alpha(\rho) - \frac{1 - a(E)}{a(E)}\varepsilon_2(E), \quad \beta_1^2(E)E_1 = \alpha(\rho) - \frac{a(E)}{1 - a(E)}\varepsilon_1(E),$$

$$\beta_2^2(E)E_2 = -\alpha(\rho) + \varepsilon_2(E), \quad \Delta\beta_1^1(E)E_1 = -1 + \frac{\varepsilon_1(E)}{1 - a(E)} \quad \text{and} \quad \Delta\beta_2^1(E)E_2 = 1 - \frac{\varepsilon_2(E)}{a(E)}.$$

Moreover, $\tau_p^1(E) = (a(E) - \alpha(\rho))/a(E)$ and $\tau_p^2(E) = (\alpha(\rho) - a(E))/(1 - a(E))$. We can also compute

$$\frac{\Delta\beta_2^1(E)}{r_2(E)} - \frac{\Delta\beta_1^1(E)}{r_1(E)} = \frac{1 - \varepsilon_1(E) - \varepsilon_2(E)}{a(E)(1 - a(E))Y(E)}.$$

Substituting in (36) yields $\xi_1/r_1 = (a - \alpha)/a + (\varepsilon_1 - (1 - a))\bar{\xi}$ with $\bar{\xi}$ as given in (63), and analogously $\xi_2/r_2 = (\alpha - a)/(1 - a) + (a - \varepsilon_2)\bar{\xi}$. The result then follows from substituting those into the adjustment formula. \square

The first two terms are simply the weighted average of the Pigouvian corrections for the two activities, since $\tau_p^1 = (a - \alpha)/a$ and $\tau_p^2 = (\alpha - a)/(1 - a)$, where the weights are the local income shares. In particular, if $a > \alpha$, meaning that activity 1 is overpaid relative to its social marginal product, then $\tau_p^1 > 0$ and $\tau_p^2 < 0$. As discussed in Section 5.3, $a\tau_p^1 + (1 - a)\tau_p^2 = 0$. In other words, the direction of no externalities, in terms of sectoral *incomes*, here is always $(a, 1 - a)$ (equivalently, in terms of sectoral effective efforts, it is $(a/r_1, (1 - a)/r_2)$, which using (62) is parallel to $(\rho, 1)$). Hence, at wage levels where the local and aggregate income shares coincide, the Pigouvian correction vanishes.

The terms in brackets capture the deviation from this weighted Pigouvian correction due to the relative return effects of a variation in the marginal tax rate. Since $\varepsilon_1 - (1 - a) = (1 - a)\Delta\beta_1^1 E_1$ and $a - \varepsilon_2 = a\Delta\beta_2^1 E_2$, this deviation can be rewritten as

$$a(1 - a) \left(\frac{f_E^1(w)}{f_E(w)} \frac{\Delta\beta_1^1 E_1}{a} + \frac{f_E^2(w)}{f_E(w)} \frac{\Delta\beta_2^1 E_2}{1 - a} \right) \bar{\xi}.$$

In particular, by the general analysis in section 4, the direction of no relative return effects, again in terms of incomes, is $(\Delta\beta_2^1 r_1, -\Delta\beta_1^1 r_2)$, which again using (62) is parallel to $(a\Delta\beta_2^1/E_1, -(1 - a)\Delta\beta_1^1/E_2)$. Hence, whenever the vector of local income shares $(f_E^1(w)/f_E(w), f_E^2(w)/f_E(w))$ points in this direction, the terms in brackets cancel, and the optimal correction coincides with the Pigouvian one.

Otherwise, the wedge depends on the sign of the relative return effects. For instance, suppose activity 1 is the overpaid one, so $a > \alpha$, and at the same time the high-wage, low redistributive preference activity. Then $\bar{\xi} > 0$ (since $I_1, R_1 > 0$ and the denominator is positive whenever the optimum involves a stable fixed point by Lemma 7 in Appendix B). Moreover, suppose that both an increase in E_1 and an increase in E_2 reduce the relative returns x^1 to activity 1 (so that $\varepsilon_1 < 1 - a$ and $\varepsilon_2 > a$). Intuitively, this would be a situation where activity 1 is, for instance, subject to crowding, whereas activity 2 mostly depresses the returns to the other activity. Then an increase in the marginal tax rate at a wage level w , by reducing both E_1 and E_2 , induces a flow of individuals into the overpaid activity 1 by increasing x^1 . Since this is not desirable, the optimal correction is in this case less than the Pigouvian correction.

A special case for the income share function $a(E)$ that leads to a particularly transparent characterization, apart from the example discussed in Section 5.3, occurs when a is homogeneous of degree zero.

C.2 Incomes Shares Homogeneous of Degree Zero

Suppose $a(E)$ is only a function of $\rho = E_1/E_2$, as is α . Then it is easy to check that $(1-a)\varepsilon_2 = a\varepsilon_1$. Hence, $\Delta\beta_1^1 E_1 = -\Delta\beta_2^1 E_2$ and the direction of no relative return effects (in E -space) is simply $(1, \rho)$, as in the preceding subsections. More importantly, as shown above, the direction of no externalities is always $(1, \rho)$ as well, so the two coincide in this case and the outer problem effectively reduces to a one-dimensional problem. This leads to a particularly simple formula for the optimal correction factor.

Corollary 2. *If $a(E)$ is homogeneous of degree zero, then*

$$\sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = \frac{1}{1-a} \left(\frac{f_E^1(w)}{f_E(w)} - a \right) \left(\frac{a-\alpha}{a} - (1-a-\varepsilon_1)\bar{\xi} \right), \quad (64)$$

where $\bar{\xi}$ is given in (63).

The first bracketed term, which parallels the corresponding terms in (39) and (42), compares the local income share from activity 1 to its aggregate income share a at each wage w . In parts of the income distribution where sector 1 dominates, the second bracketed term applies the Pigouvian correction for this sector, $\tau_p^1 = (a-\alpha)/a$, adjusted by a term that accounts for the relative return effects. These now only depend on $\varepsilon_1 = a'(\rho)\rho/a(\rho)$ since the relative return effects of E_1 and E_2 are always opposite. For instance, suppose activity 1 is again the high income activity and is overpaid, so $a > \alpha$ and $\bar{\xi} > 0$. (Note though that the labels of the sectors do not matter here; what is required to be able to sign $\bar{\xi}$ is only that the overpaid activity is also the high-income, low redistributive preference activity, so that the numerator in (63) is either positive or negative.) If $\varepsilon_1 > 1-a$, then the second bracketed term exceeds the Pigouvian correction since an increase in the marginal tax rate at w (by reducing E_1) reduces the relative return x^1 and therefore induces a beneficial shift of effort out of activity 1, and vice versa. Both the Pigouvian and relative return corrections vanish when $f_E^1(w)/f_E(w) = a$.