A Theory of Income Taxation under Multidimensional Skill Heterogeneity*

Casey Rothschild — Wellesley

Florian Scheuer — Stanford

September 2013

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 the aggregate 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: crothsch@wellesley.edu, scheuer@stanford.edu. We thank Alex Diesl for helpful comments. All errors are our own.

1 Introduction

How to design redistributive income tax systems is not only a classic question in economics, but also 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 redistributional 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

[1] See Mirrlees (1971), Diamond (1998), Saez (2001).

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.

Even though instructive as a first step, these settings remain restrictive in capturing many real-world examples. 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 productive activity cannot capture its entire social returns, because some output will be claimed by the other activity. Hence, this is a setting where both activities generate externalities, one negative and the other positive ones.

Situations with similar implications have received attention in some recent contributions to the taxation literature. 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 in this paper allows us to consider an arbitrarily large number of activities, which can be linked through arbitrarily rich externality structures: some activities may have positive, others negative externalities, and these externalities may also 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 returns to claiming credit itself, e.g. when this activity is subject to crowding. Depending on which of the two effects is stronger, the *relative* returns to the unproductive activity may rise or fall. This in turn determines whether an increase in the marginal income tax at incomes where the unproductive activity is overproportionally represented will lead to a beneficial flow of effort to the productive activity, or a perverse sectoral shift towards the *un*productive 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 framework 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 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 with demonstrating how the resulting multidimensional screening problem can be collapsed into a tractable, one-dimensional problem, extending our previous work in Rothschild and Scheuer (2013a,b) and Choné and Laroque (2010). Although settings with multidimensional heterogeneity are frequently challenging to solve (Rochet and Choné, 1998), there exists a one-dimensional, even though endogenous, summary statistic for heterogeneity in our framework. The reason is that taxes can only condition on an individual's income, not on its composition into the income shares earned through different activities. Then, for any vector of activity-specific returns, an individual always earns a given amount of income through a cost-minimizing combination of efforts in the N activities, which results in a well-defined *wage* that determines her preferences over consumption-income bundles. We can therefore reduce the screening problem to an almost standard Mirrleesian 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. The resulting fixed point conditions for these sectoral efforts show up as additional constraints in the Pareto problem, which we call consistency constraints.

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 optimal 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 for both 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, as well as various related effects induced by relative return changes (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 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 considered in our previous work 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.

While these results can all be obtained for general N-sector models, particularly transparent characterizations are available when $N = 2$. In this case, we obtain an explicit formula for the optimal corrections as a function of relative return effects and externalities from the two activities (Lemma 6). We then proceed to illustrate our results and their implications for the shape of optimal tax schedules in a number of important applications, two of which draw on our earlier work and the rest of which are novel. All of these example applications are particular cases of our general model.

The first is the externality-free case in Rothschild and Scheuer (2013b) (Proposition 4) with $N = 2$. When aggregate output has constant returns to scale and returns coincide with marginal products, the optimal adjustment factor, under general conditions, scales up marginal tax rates at the bottom of the income distribution and scales them down at the top, thus leading to a *less* progressive income tax schedule than in a standard Mirrlees model. This distortion encourages effort in the high wage and discourages effort in the low wage activity, thereby increasing the relative returns to the low wage activity under complementarity. However, this is partly counteracted by the resulting sectoral shift effects induced by this relative return change, which imply a flow of effort from the high to the low wage activity. Since this undoes some of the original increase in high-wage effort, the tax schedule is *more* progressive than in an economy without such sectoral reallocations, such as Stiglitz (1982).

The second application adds aggregate externalities in the form of increasing or decreasing returns to scale. In this case, the adjustment factor can be transparently decomposed into a local and global component (Proposition 5). The first, which depends on the relative income shares of the two activities at any given income level, is exactly the same as in the no externalities case, capturing relative return effects. The second, which is 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 examples discussed at the beginning (see Proposition 6 and its Corollaries). In this case, both the Pigouvian and relative return components of the optimal adjustment to the marginal tax rate vary across income levels. 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, because the Pigouvian tax on the high-wage activity gets weighted by an increasing income share of this activity as we move up along the income distribution. However, the optimal correction may exceed or fall short of this Pigouvian correction at any given income level depending on the relative return effects of this tax. The former case occurs 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. The first is a generalization of Rothschild and Scheuer (2013b), where all the returns depend only on the aggregate effort in one activity (Proposition 7). For instance, imagine an economy with $N-1$ traditional sectors where private and social returns are aligned, and one activity that imposes externalities on itself and all other activities. Most generally, we could allow here for positive or negative externalities, or externalities of mixed signs, imposed by this activity on all other sectors and itself, such as positive spillovers from research or entrepreneurial activities onto other sectors, but within-sector crowding effects. The last example (Proposition 8) considers the opposite case, where the returns to all activities are fixed, except for one, which depends on the aggregate efforts in all activities. In each case, we use the tools developed here to show how the optimal correction deviates from the Pigouvian correction as a function of the relative return effects.

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. Some 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 $\theta \in \Theta \equiv \Pi_{i=1}^N \Theta_i$, where the i^{th} element $\theta_i \in \Theta_i = [\underline{\theta}_i, \overline{\theta}_i]$ 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, $e = (e_1, ..., e_N)$, 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 both arguments, continuously differentiable, strictly quasiconvex and linear homogeneous.[2] We denote the consumption and vector of activity-specific efforts of an individual of type θ by $c(\theta)$ and $e(\theta) = (e_1(\theta), ..., e_N(\theta))$, 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

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

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 $E \equiv (E_1, ..., E_N)$. 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

[2] Note that this allows for preferences $\tilde{u}(c, \tilde{m}(e))$ where $\tilde{m}$ is homothetic even if not linear homogeneous: then there exists some increasing $h(.)$ 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) with linear homogeneous m. An example is $\hat{u}(c) - \sum_{i=1}^N h_i(e_i)$ when all $h_i(.)$ 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).

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 particularly simple special case would be a linear production function $Y(E) = \sum_{i=1}^{N} r_i E_i$, so individuals perform different tasks, for which they have different skills and each of which has a fixed return. A more interesting example would involve a concave constant returns to scale production function $Y(E)$ with $r_i(E) = Y_i(E)$ for all E, so returns are endogenous and respond to tax policy. For instance, in the 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 the special case $N = 2$.

Both examples above have in common that there are no externalities, i.e. private returns to all activities correspond to social marginal products. Both the recent public debate about tax policy and academic work (e.g. Rothschild and Scheuer, 2011, Piketty et al., 2013, Besley and Ghatak, 2013, Lockwood et al., 2013) have questioned this typical Mirrleesian assumption. 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 and to get credit (and compensated) for this output. If activity 1 corresponds to claiming credit and activity 2 to actual production, this could be captured in our general framework 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 the rent-seeking activity 1, and activity 1 imposes negative externalities on activity 2.

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 returns are only a (decreasing) function of effort in activity 1, whereas no other activity $i \neq 1$ imposes externalities.

Of course, our general framework here allows for much richer patterns of externalities, including positive externalities or externalities coming from or targeted in different ways at multiple other activities. Another important example would be externalities coming from increasing or decreasing returns at the aggregate level $Y(E)$, in which case effort in each activity would have positive or negative externalities, respectively, on all other activities. We will revisit all these and other applications after having developed our general theory in the next two sections.

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 $q(\theta) \equiv (q^1(\theta), ..., q^N(\theta)) \in \Delta^{N-1}$ be the vector of these income shares, where $\Delta^{N-1} = \{q \in \mathbb{R}^N | \sum_{i=1}^N q_i = 1, q_i \geq 0\}$ is the $N-1$ dimensional unit simplex.

Assuming that only income y and consumption c are observable but neither 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:

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

since each type θ can imitate another type θ' by earning the income of type θ' using a continuum of combinations of efforts and hence income shares $p = (p_1, ..., p_N)$ 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 $\{c(\theta), y(\theta), q(\theta), E\}$,*

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

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

Proof. Using (2) for $\theta = \theta'$ 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 $l(\theta) \equiv m(e(\theta)) = y(\theta) m\left(\frac{q_1(\theta)}{\theta_1 r_1(E)}, \dots, \frac{q_N(\theta)}{\theta_N r_N(E)}\right)$. □

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 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,

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

for any y. Again by linear homogeneity of m, the vector $q(\theta)$ only depends on the vector of skill ratios $\phi \equiv (\theta_1/\theta_N, ..., \theta_{N-1}/\theta_N)$ 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.

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, $-u_l(c, y/w)/\ (wu_c(c, y/w))$, 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

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

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

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

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

with the corresponding density $g_E(w, \phi)$ will be useful in the following. We denote the support of the wage distribution for any E by $[\underline{w}_E, \overline{w}_E]$, where $\underline{w}_E = w_E(\underline{\theta}_1, ..., \underline{\theta}_N)$ and $\overline{w}_E = w_E(\overline{\theta}_1, ..., \overline{\theta}_N)$. This allows us to obtain the wage 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 $w \in [\underline{w}_E, \overline{w}_E]$.[3] Finally, for any given E, we can derive Pareto weights over wages $\Psi_E(w)$, as well as their

[3] In the limiting case with $m(e) = \sum_{i=1}^N e_i$, (3) immediately implies $q_E^i(\phi) \in \{0, 1\}$ and $w_E(\theta) = \max\{\theta_1 r_1(E), ..., \theta_N r_N(E)\}$. 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.

density and sectoral decomposition $\psi_E(w) = \sum_{i=1}^N \psi_E^i(w)$, completely analogously from $\Psi(\theta)$. We are particularly interested in the regular case in which the planner assigns greater weight to low-wage individuals, i.e., where $\psi_E(w)/f_E(w)$ is non-increasing in w for any E.

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 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):

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

subject to

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

$$E_i = \frac{1}{r_i(E)} \int_{\underline{w}_E}^{\overline{w}_E} wl(w) f_E^i(w) dw \quad \forall i = 1, ..., N \tag{8}$$

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

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

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

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.[4] 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:

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

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

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

for all $w \in [\underline{w}_E, \overline{w}_E]$, 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 the uncompensated and compensated wage elasticities of total effort l as a function of the wage are $\varepsilon^u(w)$ and $\varepsilon^c(w)$, respectively.

Proof. See Appendix A.1. □

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 $\eta(w) =$

[4] See, for instance, Fudenberg and Tirole (1991), Theorems 7.2 and 7.3.

$\Psi_E(w) - F_E(w)$. 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.

The next subsections use the conditions for an optimal vector E from outer problem in order to explore the relationship between the general equilibrium Pigouvian taxes 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: the direct effect (i) 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

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

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

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

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 in E_i is therefore $l'(w) w \sum_{j=1}^N q_E^j \beta_i^j(E)$ and the average effort change at w is simply

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

where $\mathbb{E}[q_E^j(\phi)|w] = \int_\Phi q_E^j(\phi) g_E(\phi|w) d\phi$ is the average of q^j over the set $\{\theta | w_E(\theta) = w\}$ of all wage-w individuals. Defining

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

this motivates the variation in the l-schedule $\tilde{l}(w) = l(w) - l'(w) w \delta_E^i(w)$. Analogously, we can vary the V-schedule by $\tilde{V}(w) = V(w) - V'(w) w \delta_E^i(w)$. 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.[5]

We refer below to subshift 1 as the change in $l(w)$ and $V(w)$ due to the change in

[5] To wit, dropping the common argument w and using (7) and (52) yields
$\tilde{c} - c = c(\tilde{V}, \tilde{l}) - c(V, l) = \frac{1}{u_c}(\tilde{V} - V) - \frac{u_l}{u_c}(\tilde{l} - l) = \left(\frac{1}{u_c} V' - \frac{u_l}{u_c} l'\right) w \delta_E^i = \left(-\frac{u_l l}{w u_c} + \frac{u_c c' + u_l l / w}{u_c}\right) w \delta_E^i = c' w \delta_E^i.$

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

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

with

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

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 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

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

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

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

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

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

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

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

As before, we have $\sum_{j=1}^{N} I_j(E) = 0$. To interpret this, suppose $\eta(w) > 0$, so the incentive constraints bind downwards. Then a marginal increase in E_i is welfare reducing if it increases the returns to activities with $d\left(f_E^j(w)/f_E(w)\right)/dw > 0$, i.e. activities j that are locally associated with high wages in the first place, and vice versa. This is because an increase in E_i makes the wage distribution more (less) unequal in this case, which tightens (eases) 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.

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

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

It is useful to write this in terms of the Pigouvian taxes $t_p^i(E)$, $i = 1, ..., N$, 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:

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

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 $\lambda \xi_j \delta_{ij}$, 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

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

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*

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

so that we can write $\zeta_j(x_E(\phi))$ and define $\zeta \equiv (\zeta_1, .., \zeta_{N-1}, 1)$. Using this, the effective effort integrated over on the right-hand side of consistency constraint j is

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

This reveals that there are two distinct effects here: first, the change in the overall level of effort $l(w)$ 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 $l(w)$ 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

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

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

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

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

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

where

$$C_{kj}(E) \equiv \frac{1}{r_j(E)} \int_{\underline{w}_E}^{\overline{w}_E} w^2 l'(w) \text{Cov}\left(q_E^j, q_E^k \middle| w \right) f_E(w) dw \tag{26}$$

with $\text{Cov}\left(q^j, q^k\right) = \mathbb{E}\left[q^j q^k\right] - \mathbb{E}\left[q^j\right] \mathbb{E}\left[q^k\right]$. 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 $l'(w) > 0$, this variation effectively re-allocates effort towards activity j, thus increasing the right-hand side of consistency constraint j.

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

$$l(w)\theta_j \sum_{k=1}^{N} \sum_{l=1}^{N} \frac{\partial \Omega_j(\zeta(x_E(\phi)))}{\partial \zeta_l} \frac{\partial \zeta_l(x_E(\phi))}{\partial x^k} \frac{\partial x_E^k(\phi)}{\partial E_i}, \tag{27}$$

where x^l is the l-th element of the vector x. 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

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

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

$$Q_k^j(x_E(\phi)) \equiv Z_j(x_E(\phi)) \sum_{l=1}^{N} \frac{\partial \Omega_j(\zeta(x_E(\phi)))}{\partial \zeta_l} \frac{\partial \zeta_l(x_E(\phi))}{\partial x^k}. \tag{29}$$

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 relative returns x^k, but holding x fixed otherwise. Substituting this in (27), the effect becomes

$$l(w)\theta_j \sum_{k=1}^{N} \frac{w}{r_j(E)\theta_j} Q_k^j(x_E(\phi)) \frac{\partial x_E^k(\phi)}{\partial E_i} = \frac{y(w)}{r_j(E)} \sum_{k=1}^{N} \left(\beta_i^k(E) - \beta_i^N(E)\right) Q_k^j(x_E(\phi)) x_E^k(\phi)$$

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

$$-\lambda \xi_j \sum_k \left(\beta_i^k(E) - \beta_i^N(E)\right) S_{kj}(E) \tag{30}$$

with

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

We can set $S_{Nj} = 0$ for all j since $q_E^j(\phi)$ does not depend on the last element of $x_E(\phi)$. As a result, if a marginal increase in E_i increases the relative returns to activities k (i.e. $\beta_i^k(E) - \beta_i^N(E) > 0$) for which $Q_k^j > 0$ (so that an increase in x^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.

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 .

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

Proof. See Appendix A.2. □

The straightforward intuition for both sums in 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 $\int_{\underline{w}_E}^{\overline{w}_E} y(w) f_E(w) dw$, and the sectoral composition of income at wage w is irrelevant for this sum. Per (29) and the subsequent discussion, the changes $\{Q_k^j\}_{j=1,...,N}$ in $\{S_{kj}\}_{k=1,...,N}$ reflect changes in the income shares q^j, and thus do not affect $\int_{\underline{w}_E}^{\overline{w}_E} y(w) f_E(w) dw$. 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.

It is also easy to see that $\sum_{k=1}^{N} C_{kj}(E) = 0$ for all j. Intuitively, if a change in E_i does not effect any relative returns—i.e., if $\beta_i^k(E) = \beta_i^l(E)$ for all k, l—then subshift 1 causes equiproportional changes in the wages of all types θ, and hence no cross-sectoral re-allocation of effort $l(w)$ at any wage. This is useful because it means that we can equivalently write (25) as

$$- \lambda \xi_j \sum_{k=1}^{N} \left(\beta_i^k(E) - \beta_i^N(E) \right) C_{kj}(E). \tag{32}$$

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. Defining $\Delta\beta_i^j(E) \equiv \beta_i^j(E) - \beta_i^N(E)$ as the semi-elasticity of the returns $\left(\partial x_E^j(\phi)/\partial E_i\right) \Big/ x_E^j(\phi)$ establishes the following Lemma, which summarizes the results from this subsection:

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

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

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

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 $\boldsymbol{\Delta\beta}$, $\boldsymbol{C}$, and $\boldsymbol{S}$, respectively, as the matrices with $(i,j)^{\text{th}}$ 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 $\boldsymbol{I_N}$ to denote the $N \times N$ identity matrix. Then the system of optimality conditions for E_i from Lemma 4 can be written as

$$\left(\boldsymbol{I_N} - \boldsymbol{\Delta\beta}(\boldsymbol{C}+\boldsymbol{S})\right)\vec{\xi} = \vec{t}_p + \boldsymbol{\Delta\beta}\left(\vec{I}+\vec{R}\right). \tag{33}$$

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 $\sum_i \left(f_E^i(w)/f_E(w)\right)\tau_p^i(E)$, 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 $\sum_i \left(f_E^i(w)/f_E(w)\right)(\xi_i/r_i)$ from (12) includes these general equilibrium effects. We can use the system (33) 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 $\boldsymbol{\Delta\beta}$ were identically zero and there were no relative 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 $\vec{n}'\boldsymbol{\Delta\beta} = 0$, so that there are no relative wage effects in the direction $\vec{n}'$, then left-multiplying (33) 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.[6]

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 $\vec{n}'\boldsymbol{\Delta\beta} \neq 0$, 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 $\sum_i \left(f_E^i(w)/f_E(w)\right)(\xi_i/r_i) = \left(f_E^i(w)/f_E(w)\right)\tau_p^i(E) = 0$, 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, defining $\boldsymbol{\beta}$ as the matrix with (i,j)-elements $\beta_i^j(E)$:

Proposition 2. *Suppose there are at least two sectors with strictly positive earnings. Then $\vec{n}'$ is a direction of both no relative return effects and no externalities, i.e., $\vec{n}'\boldsymbol{\Delta\beta} = 0$ and $\vec{n}'\vec{t}_p = 0$, if and only if it is a left-nullvector of $\boldsymbol{\beta}$: $\vec{n}'\boldsymbol{\beta} = 0$.*

Proof. From (23), $\vec{n}'\vec{t}_p = \vec{n}'\boldsymbol{\beta}\vec{Y}$, where $\vec{Y}$ denotes the column vector of aggregate sectoral incomes $Y^i(E)$. By definition, $\boldsymbol{\Delta\beta} = \boldsymbol{\beta}(\boldsymbol{I_N} - \boldsymbol{O_N})$, where $\boldsymbol{O_N}$ 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 $\boldsymbol{I_N} - \boldsymbol{O_N}$ is zero and let $\boldsymbol{D}$ denote the matrix whose first $N-1$ columns coincide with $\boldsymbol{I_N} - \boldsymbol{O_N}$ and whose N^{th} column is $\vec{Y}$. Then $\vec{n}'\boldsymbol{\Delta\beta} = 0$ and $\vec{n}'\vec{t}_p = 0$ only if $\vec{n}'\boldsymbol{\beta D} = 0$. Since $\vec{Y} \geq 0$, with at least two strictly positive elements, $\boldsymbol{D}$ is non-singular. Hence, $\vec{n}'\boldsymbol{\beta D} = 0$ only if $\vec{n}'\boldsymbol{\beta} = 0$. □

Let $N-K$ denote the rank of the $\mathbb{R}^N \to \mathbb{R}^N$ mapping $r(E) = (r_1(E), ...r_N(E))'$ and hence of the matrix $\boldsymbol{\beta}$. 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),[7] 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.

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

[6]It is worth noting that $\boldsymbol{\Delta\beta}$ 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.

[7]Note that this implies that $r(E)$ is also sufficient for describing the variation, with E, of the tax code that leads to (33).

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 $\boldsymbol{\beta}$ is just $-\rho$ times the first row. In other words, $\boldsymbol{\beta}$ 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

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

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 the $(1,a,b)$-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 $\boldsymbol{\beta}$ is less than N.

Proposition 3. *Suppose that $\boldsymbol{\beta}$ 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.*

Proof. 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. $\boldsymbol{\beta}$ 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 a open subset of $\mathbb{R}^N$ and H from U_r onto an open subset of $\mathbb{R}^N$ such that $H(r(G^{-1}(x_1,\cdots,x_N))) = (x_1,...,x_{N-K},0,\cdots,0)$. Defining $\rho \equiv (x_1,\cdots,x_{N-K})$, we have $r(G^{-1}(\rho, x_{N-K+1},\cdots,x_N)) = H^{-1}(\rho,0,\cdots,0)$, 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, $i=1,\cdots,N-K$ is $\rho_i = G_i(\mathcal{E}(H^{-1}(\rho,0,\cdots,0);V(\cdot),l(\cdot)))$, i.e., the i^{th} component of $G(E) = G(\mathcal{E}(r(E);l(\cdot))$, written in terms of ρ. □

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 $f_E^i(w)/(f_E(w)r_i(E))$,

$\vec{\xi}$ the vector with elements ξ_i, where ξ_i are the multipliers in the original E coordinate system, and $\boldsymbol{L}$ 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, $\sum_{i=1}^{N}(f_E^i(w)/f_E(w))(\xi_i/r_i(E))$, can be replaced in the new coordinates with $\vec{n}'\boldsymbol{L}\vec{\tilde{\xi}}$, where $\vec{\tilde{\xi}}$ 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.

3.5 Eigenvalues and Stability

For the system of optimality conditions (33) to be fully informative about the optimal vector $\vec{\xi}$, the matrix $\boldsymbol{A} \equiv \boldsymbol{I_N} - \boldsymbol{\Delta\beta}(\boldsymbol{C}+\boldsymbol{S})$ needs to be invertible and hence non-singular at the optimum, which we assume in the following. If $\boldsymbol{A}$ had less than full rank, there would be multiple solutions for $\vec{\xi}$ in (33), and so the outer problem variation we used to obtain this system would not be helpful to identify $\vec{\xi}$. This assumption implies that all eigenvalues of $\boldsymbol{A}$ 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 $\boldsymbol{C}$ and $\boldsymbol{S}$ in Lemma 3 immediately implies that $\boldsymbol{A}\vec{r} = \vec{r}$, so that $\vec{r}$ is always an eigenvector of $\boldsymbol{A}$ with associated eigenvalue 1.

For the analysis of some of the examples later on, we will assume that all other eigenvalues of $\boldsymbol{A}$ are not only nonzero, but strictly positive as well. What we briefly show here is 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 (33).

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 (33) 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

$$\sum_{i=1}^{N} l^{*\prime}(w) w \delta_{E^*}^i(w)(E_i - E_i^*)$$

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.

$$l_E(w) \equiv l^*(w) - \sum_{i=1}^{N} l^{*\prime}(w) w \delta^i_{E^*}(w)(E_i - E_i^*). \tag{34}$$

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 $y_E(w) \equiv w l_E(w)$ 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

$$\mathcal{E}_i(E) = \frac{1}{r_i(E)} \int_{\underline{w}_E}^{\overline{w}_E} y_E(w) f_E^i(w) dw$$

and imagine a dynamic system with

$$\dot{E}_i = \mathcal{E}_i(E) - E_i, \quad i = 1, ..., N. \tag{35}$$

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

Lemma 5. *A fixed point E of the system (35) is stable if and only if all eigenvalues of the matrix $\boldsymbol{A} = \boldsymbol{I_N} - \boldsymbol{\Delta\beta}(\boldsymbol{C} + \boldsymbol{S})$ in (33) have positive real parts.*

4 Two Sectors

If $N = 2$, we can use the system of optimality condition (33) to solve for $\vec{\xi}$ 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

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

where

$$C(E) \equiv \int_{\underline{w}_E}^{\overline{w}_E} w^2 l'(w) \mathrm{Var}(q_E^1|w) f_E(w) dw \tag{36}$$

and

$$S(E) \equiv \int_{\underline{w}_E}^{\overline{w}_E} y(w) \int_\Phi Q_1^1(x_E^1(\phi)) x_E^1(\phi) dG_E(w,\phi). \tag{37}$$

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 $\boldsymbol{\Delta\beta}$ 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 $\boldsymbol{A} = \boldsymbol{I}_2 - \boldsymbol{\Delta\beta}(\boldsymbol{C} + \boldsymbol{S})$, i.e.,

$$\gamma_2(E) = 1 + \left(\frac{\Delta\beta_2^1(E)}{r_2(E)} - \frac{\Delta\beta_1^1(E)}{r_1(E)} \right) (C(E) + S(E)), \tag{38}$$

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 section 3.5, then $\gamma_2(E) > 0$.

Using this, solving system (33) yields:

Lemma 6. *At any Pareto optimum with $N = 2$,*

$$\vec{\xi} = \vec{t}_p + \begin{pmatrix} \Delta\beta_1^1(E) \\ \Delta\beta_2^1(E) \end{pmatrix} \frac{I_1(E) + R_1(E) + \left(\tau_p^1(E) - \tau_p^2(E)\right)(C(E) + S(E))}{\gamma_2(E)}. \tag{39}$$

Proof. See Appendix A.3. □

The system (39) makes it easy to interpret the corrective term in the marginal tax rate formula (12). As before, we obtain $\xi_i = t_p^i(E)$ 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 $\vec{n}'\vec{\xi} = \vec{n}'\vec{t}_p$. For any other $\vec{n}$, the correction term $\vec{n}'\vec{\xi}$ 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 (39) and the angle between $\vec{n}$ and the direction of no relative return effects $(\Delta\beta_2^1(E), -\Delta\beta_1^1(E))$.

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. Clearly, this is trivially the case whenever there are no externalities (so that $x = 0$, see Rothschild and Scheuer (2013b)) or only one of the two activities 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 (39) 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 7. *If $Y(E)$ is homothetic and there are no externalities then $Y(E)$ is linear homogeneous.*

Proof. Any homothetic and increasing function $Y(E)$ can be written as $Y(E) = h(\tilde{Y}(E))$ where h is increasing and $\tilde{Y}(E)$ has constant returns to scale, so that $\tilde{Y}(E) = \sum_{i=1}^N \tilde{Y}_i(E)E_i$. Since $r_i(E) = Y_i(E)$ (no externalities), $Y(E) = \sum_{i=1}^N r_i(E)E_i = \sum_{i=1}^N Y_i(E)E_i$ requires $h'(\tilde{Y}(E)) \sum_{i=1}^N \tilde{Y}_i(E)E_i = h'(\tilde{Y}(E))\tilde{Y}(E) = h(\tilde{Y}(E))$ or $h'(s)s/h(s) = 1$, i.e., h must have a constant elasticity of 1. It is easy to see that this requires h to be of the form $h(s) = cs$, where c is some constant: Integrate both sides of $h'(s)/h(s) = 1/s$ to get $\log(h(s)) = \log(s) + \text{const.}$, or $h(s) = cs$. Hence, $Y(E)$ itself has constant returns to scale. □

In the two-sector case, the returns $r_i(E) = Y_i(E)$ are homogeneous of degree zero, and thus only depend on $\rho \equiv E_1/E_2$. 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 and applying Lemma 6, we obtain the following characterization as in Rothschild and Scheuer (2013b):

Proposition 4. *If $N = 2$, $Y(E)$ is homothetic and there are no externalities, then the numerator in the marginal tax rate formula (12) is*

$$1 - \sum_{i=1}^{2} \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i(E)} = 1 + \left(\frac{f_E^1(w)}{f_E(w)} - \alpha(\rho) \right) \bar{\xi} \;\; with \;\; \bar{\xi} \equiv \frac{(I_1 + R_1)/\sigma}{\alpha(1-\alpha)Y + (C+S)/\sigma}. \quad (40)$$

Proof. Straightforward calculations yield $\beta_1^1(E) = -\beta_2^1(E)/\rho = Y_1'(\rho)/(E_2 Y_1(\rho))$, $\beta_1^2(E) = -\beta_2^2(E)/\rho = Y_2'(\rho)/(E_2 Y_2(\rho))$, $\Delta\beta_1^1(E) = -1/(E_1\sigma(\rho))$, $\Delta\beta_2^1(E) = 1/(E_2\sigma(\rho))$ where $-1/\sigma(\rho) = \rho(Y_1'(\rho)/Y_1(\rho) - Y_2'(\rho)/Y_2(\rho))$, and $t_p^1(E) = t_p^2(E) = 0$. Substituting in (39) yields $\xi_1/r_1(E) = -(1-\alpha(\rho))\overline{\xi}$ and $\xi_2/r_2(E) = \alpha(\rho)\overline{\xi}$, and substituting both in the adjustment term delivers the result. □

This makes clear that the adjustment factor vanishes when technology becomes linear, so $\sigma(\rho) \to \infty$ and $\overline{\xi} \to 0$. 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 $R_1(E) > 0$. Then $\overline{\xi} > 0$, which means that 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.

In particular, the top marginal tax rate is $T'(y(w)) = \left(\alpha - f_E^1(w)/f_E(w)\right)\overline{\xi} < 0$. As discussed in Rothschild and Scheuer (2013b), 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.[8]

Finally, the adjustment vanishes precisely 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, so there are no Pigouvian motives for taxation and any nonzero adjustment term in (12) is exclusively due to relative return effects.

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

[8]Note 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.

$\tilde{Y}$-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,

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

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 6 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 given by*

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

where $\overline{\xi}$ is given in (40).

Proof. 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 (39) yields $\xi_1/r_1 = 1 - \varepsilon_h(E) - (1-\alpha(\rho))\overline{\xi}$ and $\xi_2/r_2 = 1 - \varepsilon_h(E) + \alpha(\rho)\overline{\xi}$, and using this in (12) yields (42). □

Unsurprisingly, (42) collapses to (40) if $\varepsilon_h = 1$ and we are back to a setting with constant returns to scale and no externalities. Otherwise, the optimal adjustment in (42) can be transparently decomposed into two parts: the first, local one (which varies with the wage and hence income level and therefore affects the progressivity of the optimal tax schedule) is exactly the same as in Proposition 4. The second, new component is $1 - \varepsilon_h(E)$ and 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 returns to scale and so marginal tax rates are scaled up relative to an economy with constant returns to scale, whereas they are scaled down if $\varepsilon_h(E) > 1$ (so there are globally positive externalities from increasing 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.

Observe that 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 properties 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 income 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 obviously points into a different quadrant.[9] Hence, $\boldsymbol{\beta}$ has full rank and both consistency constraints are needed in this example.

5.3 General Sectoral Income Shares

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: total output $Y(E)$ 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 composition $a(E)$ and $1 - a(E)$ across the two sectors. If $a(E) \neq \alpha(\rho)$, effort in one of the two activities is underpaid relative to its marginal product, and the other is overpaid.

An extreme example would be $Y(E) = E_2$, 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, increasing incomes at the expense of others without adding anything to aggregate resources (so it is overpaid for any $a > 0$). Our general formulation allows for all intermediate cases where effort in an activity may both contribute to output and increase its income share, and the latter possibly decoupled from social marginal products.

Formally, let $Y^1(E) = a(E)Y(E)$ and $Y^2(E) = (1 - a(E))Y(E)$, so that, by linear returns,

$$r_1(E) = a(E)Y(E)/E_1 \quad \text{and} \quad r_2(E) = (1 - a(E))Y(E)/E_2. \tag{43}$$

[9] Moreover, 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 $\varepsilon_h = 1$.

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 6. *If $N = 2$, $Y(E)$ has constant returns to scale and private returns are given by (43), then the adjustment to the marginal tax rate formula (12) is*

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

with

$$\overline{\xi} = \frac{I_1 + R_1 + \frac{a-\alpha}{a(1-a)}(C+S)}{a(1-a)Y + (1-\varepsilon_1-\varepsilon_2)(C+S)}. \tag{44}$$

Proof. We have

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

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

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 (39) yields $\xi_1/r_1 = (a-\alpha)/a + (\varepsilon_1 - (1-a))\overline{\xi}$ with $\overline{\xi}$ as given in (44), and analogously $\xi_2/r_2 = (\alpha - a)/(1-a) + (a-\varepsilon_2)\overline{\xi}$. The result then follows from substituting those into the adjustment formula. □

This formula is intuitive. 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 $\tau_p^2 < 0$. Because the externalities here only result from the distribution of incomes across sectors, not from overall output (which was the case in the preceding section), the Pigouvian correction is zero at the aggregate level: $a\tau_p^1 + (1-a)\tau_p^2 = 0$. 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 (43) 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

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

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 (43) 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 $\overline{\xi} > 0$ (since $I_1, R_1 > 0$ and the denominator is positive whenever the optimum involves a stable fixed point by Lemma 5). 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.

Two special cases for the income share function $a(E)$ are of separate interest and lead to particularly transparent characterizations: when a is homogeneous of degree zero and the example from the beginning, where a only depends on E_1 and $Y = E_2$.

5.3.1 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 1. *If $a(E)$ is homogeneous of degree zero, then*

$$\sum_{i=1}^{2} \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = \frac{1}{1-a} \left(\frac{f_E^1(w)}{f_E(w)} - a \right) \left(\frac{a-\alpha}{a} - (1-a-\varepsilon_1)\overline{\xi} \right), \tag{45}$$

where $\overline{\xi}$ is given in (44).

The first bracketed term, which parallels the corresponding terms in (40) 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 $\overline{\xi} > 0$.[10] 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 $f_E^1(w)/f_E(w) = a$.

It is also worth pointing out that (45) has sharp implications for the optimal progressivity of the income tax schedule. In particular, under the conditions in the previous paragraph, the adjustment term in (45) is positive at high income levels and negative otherwise. In other words, 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 emphasized here then determine whether the progressivity is even more or less pronounced than in this Pigouvian benchmark.

5.3.2 A Pure Resource Transfer Activity

Let us return to the extreme example where $a(E) = a(E_1)$ and $Y(E) = E_2$, so that only activity 2 is productive, whereas activity 1 only means to capture resources produced by others. 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

[10]Note that the labels of the sectors do not matter here; what is required to be able to sign $\overline{\xi}$ is only that the overpaid activity is also the high-income, low redistributive preference activity, so that the numerator in (44) is either positive or negative.

Pigouvian subsidy.

Applying Proposition 6 with $\alpha = \varepsilon_2 = 0$ yields

Corollary 2. *If $a(E) = a(E_1)$ and $Y(E) = Y(E_2)$, then*

$$\sum_{i=1}^{2} \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = \left(1 - (1 - a - \varepsilon_1)\overline{\xi}\right) \frac{f_E^1(w)}{f_E(w)} + \left(-\frac{a}{1-a} + a\overline{\xi}\right) \frac{f_E^2(w)}{f_E(w)} \tag{46}$$

with

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

Again, $\overline{\xi} \geq 0$ 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 $\boldsymbol{A}$, so it is positive in a stable fixed point according to Lemma 5).

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 $\varepsilon_1 \gtrless 1 - a$ as before. 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 (46) closer to (45) is

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

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. As before, 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 $a/(1-\varepsilon_1)$.

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 $i=1,...,N$. 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 $i=2,...,N$ and all j, the matrix $\boldsymbol{\beta}$ has rank one (all rows except for the first are zero) and the optimality condition (33) 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. In fact, we can use (33) to explicitly solve for ξ_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 $t_p^1=-\sum_{j=1}^{N}\beta_1^jY^j$ and C_{j1} and S_{j1} are given in (26) and (31). This leads to the following result:

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

$$\bar{\xi} = \frac{\tau_p^1 + \sum_{j=1}^{N-1} \Delta\beta_1^j (I_j + R_j)/r_1}{1 - \sum_{j=1}^{N-1} \Delta\beta_1^j (C_{j1} + S_{j1})}. \tag{47}$$

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 (47) 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.[11]

The second term in the numerator of (47) 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).[12] Then an increase in the marginal income tax at wage levels where activity 1 is prevalent raises the returns to the the lower wage, high redistributive preference activities by decreasing E_1, thus achieving indirect redistribution. Of course, analogous results can be obtained from (47) 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 (36) and (37). 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

[11]Note that the denominator of (47) is always positive when the optimum involves a stable fixed point for E_1 in terms of Lemma 5, since it is the eigenvalue of the matrix $\boldsymbol{A}$ associated with the unit eigenvector $(1, 0, 0, ..., 0)'$.

[12]Recall 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.

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 $\overline{\xi} > \tau_p^1$ 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.

Finally, this special formula would also result in the general N-sector model whenever $\Delta\beta_1^j = 0$ for all $j \neq 1$, i.e. if changes in E_1 only affect the relative return to activity 1 itself, but not the relative returns to the externality receiving activities $j = 2, ..., N$. A simple example would involve returns $r_1(E_1)$ and $r_j(E_1) = \bar{r}_j r_N(E_1)$ for all $j = 2, ..., N-1$. Then variations in E_1 do not induce any shifts among sectors $2, ..., N$, but only between them and sector 1, so the model effectively collapses to a two-sector setting as well.

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 $i = 2, ..., N$. A simple example with $N = 2$ would be another specification of the economy where one activity is just capturing output produced by others, 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.[13]

Generally, $\beta_i^j = 0$ for all $j \neq 1$ and $\boldsymbol{\beta}$ 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 $\boldsymbol{\Delta\beta} = \boldsymbol{\beta}$). 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 $(\beta_1^1, \beta_2^1, ..., \beta_N^1)$ 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, (33) immediately implies $\xi_i/\beta_i^1 = \xi_1/\beta_1^1$ for all i and therefore the following result:

[13]For $N = 2$, this is a special case of the example considered in subsection 5.3 with $a(E) = 1 - E_2/Y(E_1)$. But we could imagine a more general model with $N-1$ stealing sectors $i = 2, ..., N$ 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*

$$\sum_{i=1}^{N} \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = \sum_{i=1}^{N} \frac{f_E^i(w)}{f_E(w)} \frac{\beta_i^1}{r_i} \bar{\xi} \quad \text{with} \quad \bar{\xi} = \frac{-Y^1 + I_1 + R_1}{1 - \sum_{i=1}^{N} \beta_i^1 (C_{1i} + S_{1i})}. \tag{48}$$

Proof. Using $\xi_i = \xi_1 \beta_i^1 / \beta_1^1$, we can use the first equation in the system (33) to solve for ξ_1:

$$\xi_1 = \left(t_p^1 + \beta_1^1 (I_1 + R_1)\right) \Big/ \left(1 - \sum_{i=1}^{N} \beta_i^1 \left(C_{1i} + S_{1i}\right)\right)$$

Again using $\xi_i = \xi_1 \beta_i^1 / \beta_1^1$ and $t_p^i = t_p^1 \beta_i^1 / \beta_1^1$ delivers the result. □

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 $\beta_i^1(E) > 0$ 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 $\bar{\xi}$ in (48) 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.[14]

The term $I_1 + R_1$ in the numerator of $\bar{\xi}$ 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

[14]Note that the denominator of $\bar{\xi}$ in (48) is positive if the optimum involves a stable fixed point of E according to Lemma 5, since it is an eigenvalue of the matrix $\boldsymbol{A}$ associated with the eigenvector $(\beta_1^1, \beta_2^1, ..., \beta_N^1)'$.

reduces to $1 - (\beta_1^1/r_1 - \beta_2^1/r_2)(C + S)$ with C and S given in (36) and (37). 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 in r_1 otherwise. The denominator scales the direct effects in the numerator to account for these indirect multiplier effects of the sectoral shifts on r_1.

A particularly simple adjustment term results if β_i^1 is proportional to r_i for all i, so $\beta_i^1 = x r_i$ for some x. In this case, the adjustment term becomes

$$\sum_{i=1}^{N} \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = x(-Y^1 + I_1 + R_1)$$

by Lemma 3, i.e. it reduces to just the direct effects of the marginal tax. Note that this is independent of w and therefore a uniform adjustment to the tax schedule. Intuitively, a change in each sectoral income Y^i has the same effect on r_1 in this case, so the optimal correction for each activity is just x (there are no indirect effects from sectoral reallocations) and the local sectoral income shares become irrelevant.

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 (33) 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 5.3.2, 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*, forthcoming.

BOOTHBY, W. (1986): *An Introduction to Differentiable Manifolds and Reimannian Geometry*. Academic Press, Orlando, Florida, 2nd edn.

CHONÉ, P., AND G. LAROQUE (2010): "Negative Marginal Tax Rates and Heterogeneity," *American Economic Review*, 100, 2532–2547.

DIAMOND, P. (1998): "Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Tax Rates," *American Economic Review*, 88, 83–95.

FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. MIT Press, Cambridge.

LOCKWOOD, B., C. NATHANSON, AND G. WEYL (2013): "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.

SAEZ, E. (2001): “Using Elasticities to Derive Optimal Tax Rates,” *Review of Economic Studies*, 68, 205–229.

STIGLITZ, J. (1982): “Self-Selection and Pareto Efficient Taxation,” *Journal of Public Economics*, 17, 213–240.

A Proofs for Sections 3 and 4

A.1 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),

$$\begin{aligned}\mathcal{L} \quad = \quad & \int_{\underline{w}_E}^{\overline{w}_E} V(w)\psi_E(w)dw - \int_{\underline{w}_E}^{\overline{w}_E} V(w)\hat{\eta}'(w)\lambda dw + \int_{\underline{w}_E}^{\overline{w}_E} u_l(c(V(w),l(w)),l(w))\frac{l(w)}{w}\hat{\eta}(w)\lambda dw \\ & + \sum_{i=1}^{N}\xi_i\lambda\left[E_i - \frac{1}{r_i(E)}\int_{\underline{w}_E}^{\overline{w}_E} wl(w)f_E^i(w)dw\right] + \lambda\int_{\underline{w}_E}^{\overline{w}_E}(wl(w) - c(V(w),l(w)))\,f_E(w)dw. \end{aligned} \tag{49}$$

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

$$\hat{\eta}'(w)\lambda = \psi_E(w) - \lambda f_E(w)\frac{1}{u_l(w)} + \hat{\eta}(w)\lambda\frac{u_{cl}(w)}{u_l(w)}\frac{l(w)}{w}. \tag{50}$$

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

$$\eta'(w) = \psi_E(w)\frac{u_c(w)}{\lambda} - f_E(w) + \eta(w)\frac{u_{cc}(w)c'(w) + u_{cl}(w)l'(w) + u_{cl}(w)l(w)/w}{u_c(w)}. \tag{51}$$

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

$$u_c(w)c'(w) + u_l(w)l'(w) + u_l(w)\frac{l(w)}{w} = 0, \tag{52}$$

the fraction in (51) 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 (51) and rearranging yields

$$-\frac{\partial MRS(w)}{\partial c}l(w)\frac{y'(w)}{y(w)}\eta(w) = f_E(w) - \psi_E(w)\frac{u_c(w)}{\lambda} + \eta'(w). \tag{53}$$

Integrating this ODE gives

$$\begin{aligned} \eta(w) &= \int_w^{\overline{w}_E} \left(f_E(s) - \psi_E(s)\frac{u_c(s)}{\lambda}\right) \exp\left(\int_w^s \frac{\partial MRS(t)}{\partial c} l(t)\frac{y'(t)}{y(t)}dt\right) ds \\ &= \int_w^{\overline{w}_E} \left(1 - \frac{\psi_E(s)}{f_E(s)}\frac{u_c(s)}{\lambda}\right) \exp\left(\int_w^s \left(1 - \frac{\varepsilon^u(t)}{\varepsilon^c(t)}\right)\frac{dy(t)}{y(t)}\right) f_E(s)ds, \end{aligned} \tag{54}$$

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

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

which after some algebra can be rewritten as

$$w f_E(w)\left(1 - \frac{MRS(w)}{w}\right) - w\sum_{i=1}^N \frac{\xi_i}{r_i(E)} f_E^i(w) = \eta(w)\left(\frac{\partial MRS(w)}{\partial l}\frac{l}{w} + \frac{MRS(w)}{w}\right), \tag{55}$$

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

$$1 - \sum_{i=1}^N \frac{\xi_i}{r_i(E)}\frac{f_E^i(w)}{f_E(w)} = (1 - T'(y(w)))\left[1 + \frac{\eta(w)}{w f_E(w)}\left(1 + \frac{\partial MRS(w)}{\partial l}\frac{l}{MRS(w)}\right)\right]. \tag{56}$$

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 (54) and (56).

A.2 Proof of Lemma 3

For C_{kj}, this follows from

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

because $\sum_{j=1}^{N} \text{Cov}\left(q_E^j, q_E^k \middle| w\right) = \text{Cov}\left(\sum_{j=1}^{N} q_E^j, q_E^k \middle| w\right) = \text{Cov}\left(1, q_E^k \middle| w\right) = 0$ 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 $\phi \in \Phi$. To see this, we use (28) and (29) to write

$$0 = \frac{\partial 1}{\partial x^k} = \sum_{j=1}^{N} \frac{\partial q_E^j(\phi)}{\partial x^k} = \sum_{j=1}^{N} \frac{\partial Z_j(x_E(\phi))}{\partial x^k} \Omega_j(\zeta(x_E(\phi))) + \sum_{j=1}^{N} Q_k^j(x_E(\phi)).$$

Hence, the result is established if we show that $\sum_j \Omega_j \partial Z_j / \partial x^k = 0$. Using (28), we have

$$\frac{\partial Z_j}{\partial x^k} = \delta_{kj} m - x^j m_k \frac{q^k}{\left(x^k\right)^2}, \tag{57}$$

where we suppressed the arguments of m and m_k denotes the partial derivative of m w.r.t. its k-th argument. Note that the first order conditions for the minimization in (28) are $m_k / x^k = m_N$ for all $k = 1, ..., N-1$, which implies

$$\sum_{k=1}^{N-1} m_k \frac{q^k}{x^k} = m_N \sum_{k=1}^{N-1} q^k = m_N(1 - q^N). \tag{58}$$

On the other hand, by Euler's theorem and linear homogeneity of m, we have

$$\sum_{k=1}^{N-1} m_k \frac{q^k}{x^k} + m_N q^N = m. \tag{59}$$

Combining (58) and (59) implies $m_N = m$ and hence $m_k / x^k = m$ for all $i = k, ..., N-1$. Substituting this in (57) yields $x^k \partial Z_j / \partial x^k = m\left(\delta_{kj} x^k - x^j q^k\right) = m x^j \left(\delta_{kj} - q^k\right)$ since $\delta_{kj} x^k = \delta_{kj} x^j$. Using this and (24), we have

$$\sum_{j=1}^{N} \Omega_j x^k \frac{\partial Z_j}{\partial x^k} = \sum_{j=1}^{N} \zeta_j x^j \left(\delta_{kj} - q^k\right) = \sum_{j=1}^{N} \frac{q^j}{q^N} \left(\delta_{kj} - q^k\right) = \frac{1}{q^N} \sum_{j=1}^{N} \left(\delta_{kj} q^j - q^j q^k\right) = \frac{1}{q^N} \left(q^k - q^k\right) = 0.$$

Dividing through by x^k yields the desired $\sum_j \Omega_j \partial Z_j / \partial x^k = 0$, which establishes the result.

A.3 Proof of Lemma 6

Dropping the arguments E, the optimality conditions (33) can be written for $N = 2$ as

$$\boldsymbol{A} \vec{\xi} = \vec{t}_p + \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} (I_1 + R_1). \tag{60}$$

Since $\left(\Delta\beta_1^1, \Delta\beta_2^1\right)'$ is an eigenvector of $\boldsymbol{A}$, it is also an eigenvector of $\boldsymbol{A}^{-1}$ (with associated eigenvalue $1/\gamma_2$), and we can write (60) as

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

Moreover, defining the eigenbasis

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

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

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

Hence,

$$\vec{\xi} = \vec{t}_p + \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix}\frac{I_1 + R_1 - b\left(\Delta\beta_2^1/r_2 - \Delta\beta_1^1/r_1\right)(C+S)}{\gamma_2}. \tag{61}$$

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

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

Substituting in (61) yields the result.