

A Theory of Income Taxation under Multidimensional Skill Heterogeneity*

Casey Rothschild
Wellesley

Florian Scheuer
Stanford and NBER

December 2014

Abstract

We develop a unifying framework for optimal income taxation in multi-activity economies with general production technologies. Agents are characterized by an N -dimensional skill vector that captures intrinsic abilities in N activities. The private return to each activity depends on individual skill and an aggregate activity-specific return, which is a general function of the economy-wide distribution of efforts across activities. The optimal tax schedule features a multiplicative income-specific correction to an otherwise standard tax formula. Because taxes affect the relative returns to different activities, this correction diverges, in general, from the weighted average of the Pigouvian taxes that would align private and social returns in each activity. We characterize this divergence as a function of relative return elasticities, and its implications for the shape of the income tax both generally and in a number of applications, including externality-free economies with general equilibrium effects, economies with increasing or 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 Spencer Bastani, Lizi Chen, Gordon Dahl, Alex Diesl, Roger Gordon, Laurence Jacquet, Guido Menzio, Torsten Persson, Edouard Schaal, Chris Sleet, Johannes Spinnewijn, Aleh Tsyvinski, Venky Venkateswaran, Gianluca Violante and seminar participants at Bocconi, Georgia State, Harvard, IIES Stockholm, LSE/IFS, NYU, University of Pennsylvania, Toulouse, UC San Diego, Wellesley, Zurich, the 2014 Taxation Theory Conference (Cologne), Barcelona GSE Summer Forum on Macro and Micro Perspectives on Taxation and National Tax Association Annual Meeting (Santa Fe) for helpful comments. Scheuer and Rothschild respectively thank Harvard University and the Radcliffe Institute for Advanced Study at Harvard University for funding and hospitality. All errors are our own.

1 Introduction

How to design redistributive income tax systems is both a classic question in economics and a recurrent topic in public policy debates, as exemplified by the recent “Occupy” and “Tea Party” movements. While the standard equity-efficiency tradeoff, i.e., the tension between redistributive goals and tax distortions, which has long been emphasized by the formal optimal taxation literature,¹ has played some role, the recent debate has pointed to two central issues largely absent from this canonical framework. First, the trend towards greater income inequality in the past decades (as documented e.g. by Atkinson, Piketty and Saez, 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, some supporters of 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 framework for the analysis of optimal income taxation in multi-activity economies with fully general production technologies. In particular, individuals can pursue N different activities, the returns to each of which depend in a fully general way on the aggregate efforts in all N activities (and not necessarily aligned with marginal products). Naturally, we allow for N -dimensional heterogeneity of privately known individual skills across all N activities. Tax policies in this setting reflect two novel effects: First, across-activity shifts of effort caused by income-tax-change-induced changes in the relative returns to different activities; 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 (2013), we 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 (2014), we 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 exceed its social marginal product, and the second, traditional activity generates no externalities. We show that general equilibrium effects from sectoral shifts of effort between productive and unproductive work cause the optimal income-tax correction to diverge from the

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

partial equilibrium Pigouvian externality correction.

These instructive examples remain restrictive for capturing many real-world settings. Imagine, as a stylized example, a team production setting where individuals spend effort both to actually produce output and also to claim credit (and get paid) for output they or others have produced. Claiming credit is a zero sum activity from a social perspective, so its private returns exceed its social returns. On the other hand, part of the productive activity's returns are captured by credit-claiming effort. Hence, this is a setting where both activities generate externalities—one negative, and the other positive.

Some recent contributions to the taxation literature have addressed related phenomena. For instance, Ales, Kurnaz, and Sleet (2014) examine the effect of technical change on the optimal income tax in a multi-sector economy that is a special case of ours, abstracting from externalities and multidimensional heterogeneity. Piketty, Saez, and Stantcheva (2014) emphasize that some top incomes may come at the expense of lower incomes, e.g. because CEOs 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 others, e.g. in the form of bailouts in the financial sector financed by taxes on everyone else. Lockwood, Nathanson, and Weyl (2014) 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 the papers incorporating externalities assume a particular pattern thereof, where whenever some activity is overpaid, this comes at the expense of everyone else *uniformly*, rather than potentially at the expense of some more than others.

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

These activity shifts in response to relative return changes turn out to play an important role for 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 at a given income level induces no relative return changes. We use these general results in various applications to characterize the optimal progressivity of the income tax schedule for any redistributive objectives, captured by arbitrary Pareto weights.

Since our model naturally involves N dimensions of private information, we begin by demonstrating how they can be collapsed into a single dimension relevant for screening, extending our previous work in Rothschild and Scheuer (2013, 2014).² In particular, we identify a one-dimensional, but endogenous, summary statistic for heterogeneity in our framework. When confronted with an income tax, an individual always earns a given amount of income through a cost-minimizing combination of efforts in the N activities. For any vector of activity-specific returns, this results in a well-defined *wage* that determines her preferences over consumption-income bundles. We can therefore work with a screening problem in terms of these wages, with the only complication that they depend on sectoral returns and therefore the vector of aggregate efforts in all activities.³

We first solve this screening problem for any given combination of aggregate activity-specific efforts (the “inner” problem). We obtain a formula for the Pareto optimal marginal income tax rates (Proposition 1) which closely mirrors the standard Mirrlees formula but which features an additional adjustment factor capturing the optimal corrections for both externalities and relative return effects. The remainder of the paper then focuses on precisely characterizing this adjustment factor. This characterization is closely related to the “outer” problem of finding the optimal combination of aggregate efforts in each activity (for given Pareto weights), and we describe in detail the welfare effects of marginal variations in these efforts.

To interpret the adjustment factor in the marginal tax rate formula, we 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. 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

²Other recent taxation studies under multidimensional heterogeneity include Kleven, Kreiner and Saez (2009), Choné and Laroque (2010), Jacquet, van der Linden and Lehmann (2013), Scheuer (2014), Jacquet and Lehmann (2014), Gomes, Lozachmeur, and Pavan (2014), and Golosov, Tsyvinski, and Werquin (2014).

³Multidimensional heterogeneity therefore only has non-trivial effects in our framework when there are general equilibrium effects. Otherwise, a standard tax formula applies, as in the multidimensional screening settings with linear technology considered in Jacquet and Lehmann (2014) and Hendren (2014).

effort vectors in which there are neither relative return effects nor externalities, then the outer problem collapses to an $N - K$ -dimensional problem with $N - K$ consistency constraints. The special cases in Rothschild and Scheuer (2013, 2014), where two-sector models can be solved with a single consistency constraint, are applications of this principle.

We then illustrate how the tools we develop here can be used to characterize optimal tax schedules in several important applications, two of which extend our earlier work and the rest of which are novel. First, we investigate how the results from the externality-free environment with two sectors in Rothschild and Scheuer (2013) extend to more than two sectors (Proposition 4) and show that the additional sectors can reinforce the regressive adjustment to the standard Mirrleesian tax schedule, effectively moving the optimal income tax closer to that in a model with fixed occupations, such as Stiglitz (1982).⁴

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

We then consider the case where aggregate technology exhibits constant returns to scale but sectoral income shares are decoupled from marginal products, as motivated by the credit-claiming example discussed above (Proposition 6). For instance, suppose the relatively high-wage activity is overpaid, in the sense that its aggregate income share exceeds what would correspond to its marginal product. Then the Pigouvian correction implies a *more* progressive income tax schedule than in a standard Mirrlees model. However, since the optimal correction (e.g.) exceeds the Pigouvian correction when 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—the optimal income tax schedule may be *even more* progressive than under the Pigouvian benchmark.

Finally, we turn to two applications that we can fully characterize for general N , namely the case where all returns depend only on the aggregate effort in one activity (Proposition 7), and the case where the returns to all activities are fixed, except for one, which depends on the aggregate efforts in all activities (Proposition 8). The first is a generalization of Rothschild and Scheuer (2014), allowing for more than one traditional

⁴Naito (1999) examines commodity taxation and production efficiency in Stiglitz’s (1982) two-class economy. Gomes, Lozachmeur, and Pavan (2014) and Scheuer (2014) consider differential taxation of the sectors and its implication for production efficiency, in contrast to our focus on optimal income taxation.

activity and positive or mixed externalities, such as positive spillovers from research or entrepreneurial activities onto other sectors, but within-sector crowding effects.

Our paper is part of the growing body of work discussed above that addresses the design of the income tax in multi-activity economies. As a result, our approach crucially differs from most of the literature on corrective taxation in the presence of externalities.⁵ The simple “principle of targeting” (Dixit, 1985) does not apply in our setting because we rule out fine-tuned instruments that perfectly discriminate among specific activities. This makes our analysis both theoretically interesting and practically relevant. We show how, under imperfect instruments, the optimal correction diverges from the Pigouvian tax, which would apply under perfect targeting, depending on relative return effects.

Our analysis is more related to Diamond (1973), although our motivation, framework, and instruments are quite distinct. Most importantly, Diamond considers linear commodity taxes in the Ramsey tradition, while we work in a Mirrleesian, non-linear income tax setting, which combines redistributive and corrective motives for taxation. He shows that the optimal linear tax on an externality producing consumption good can be decomposed into a term that captures the direct effect of the tax on the demand for the good, and another term that reflects the indirect effect of the changes in consumption of the good induced by the direct effect. Our general equilibrium effects are very different, as they result from effort choice along N intensive margins. Moreover, we are able to characterize in which direction and by how much the optimal correction should deviate from the Pigouvian tax rate as a function of simple properties of technology that could potentially be estimated empirically (see e.g. Ales, Kurnaz and Sleet, 2014).

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

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 the N -vector $\theta \in \Theta \equiv \prod_{i=1}^N \Theta_i$ of unobservable skills

⁵See Sandmo (1975), Sadka (1978), Cremer, Gahvari, and Ladoux (1998), and Kopczuk (2003).

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

Individual preferences are described by a continuously differentiable and concave utility function over consumption c and the vector of efforts in each activity, $e = (e_1, \dots, e_N)$, given by $U(c, e) = u(c, m(e)) \equiv u(c, l)$. We assume $u_c > 0$, $u_l < 0$, and that the effort aggregator $m(e)$ is increasing in all arguments, continuously differentiable, strictly quasi-convex and homogenous of degree 1.⁶ We denote the consumption and vector of activity-specific efforts of an individual of type θ by $c(\theta)$ and $e(\theta) = (e_1(\theta), \dots, e_N(\theta))$, and the total individual effort and utility by $l(\theta) \equiv m(e(\theta))$ and $V(\theta) \equiv u(c(\theta), l(\theta))$.

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

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

is the aggregate effective (i.e., skill-weighted) effort in activity i , and each Y_i can depend on the entire vector of aggregate efforts $E \equiv (E_1, \dots, E_N)$. The income of an individual of type θ attributable to 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 .

Since Y and Y_i are arbitrary functions of E , our framework employs what is, to our knowledge, the most general production technology considered heretofore in the Mirrleesian taxation literature. Our only substantive assumption is that each unit of effective effort *within* a given activity is equally remunerated.⁷ 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

$$Y(E) = \sum_{i=1}^N Y_i(E) = \sum_{i=1}^N r_i(E) E_i.$$

The returns r_i may deviate from $\partial Y(E) / \partial E_i$, i.e., the marginal product of effort in activity i , which allows for interesting linkages across sectors, as we now discuss.

2.2 Examples

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

⁶Redefining $u(c, l) \equiv \tilde{u}(c, h(l))$ allows for preferences $\tilde{u}(c, \tilde{m}(e))$ with *any* increasing and homothetic \tilde{m} (equal to $h(m(e))$, for some increasing $h(\cdot)$ and linear homogeneous $m(e)$). A limiting case is $m(e) = \sum_{i=1}^N e_i$, wherein individuals specialize in their highest-return activity, as in Rothschild and Scheuer (2013).

⁷With arbitrary N , this is itself unrestrictive: activities with imperfectly substitutable effective effort can be sub-divided until the assumption is satisfied.

in the limiting case where $m(e)$ becomes linear and individuals always specialize in one of the N activities, $Y(E)$ is a production function for a Roy (1951) model economy with N complementary sectors or occupations i , as in Rothschild and Scheuer (2013) for $N = 2$. Ales, Kurnaz and Sleet (2014) simulate optimal income taxes for a similar economy in which aggregate efforts in the activities (corresponding to occupations) are combined through CES technology

$$Y(E) = A \left[\sum_{i=1}^N b_i E_i^{\frac{\epsilon-1}{\epsilon}} \right]^{\frac{\epsilon}{\epsilon-1}}$$

but skill heterogeneity is effectively reduced to a single dimension.⁸

When private returns coincide with social marginal products in all activities as in these examples, technology must exhibit constant returns to scale (by Euler's theorem). Our general framework also allows us to consider tax policy when returns and marginal products are not aligned. A misalignment can arise first if Y exhibits non-constant returns to scale, which implies positive or negative aggregate effort externalities.

Second, as emphasized by the recent policy debate, externalities can arise, even with constant returns to scale, when some activities are over- or under-compensated relative to their marginal product. For example, consider again the team production setting from the introduction where individuals exert effort both to produce output (activity 2) and to get credit (and compensated) for this output (activity 1). This can be captured by $Y(E) = E_2$ and $Y_1(E) = a(E_1)E_2$, $Y_2(E) = (1 - a(E_1))E_2$, where $a(E_1)$ is some increasing function. Here, activity 2 generates positive externalities as it increases the returns $r_1 = a(E_1)E_2/E_1$ to activity 1, and activity 1 imposes negative externalities on activity 2. For instance, in Biais, Foucault and Moinas (2011), fast traders impose externalities on slow traders through adverse selection from their information advantage. In Glode and Lowery (2012), financial sector workers engage in both (unproductive) speculative trading and surplus creation (e.g. from market making) with interlinked profits from both activities.

Another example for a pure zero-sum activity is a setting where activity 1 just takes away output produced in activity 2 one-for-one (e.g. through bargaining), 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 (2014) (again for $N = 2$), where only one (rent-seeking) activity imposes (negative) externalities on itself and all other activities, so

⁸In particular, there is an interval of types $k \in [0, 1]$ such that k 's skill in activity i is $\theta_k(i)$ and higher types $k' > k$ have both absolute and relative advantage in higher activities $i' > i$. This specifies a one-dimensional curve in our N -dimensional skill space, ruling out overlapping wage distributions in the occupations.

$r_i(E_1)$ for all i and all r_i are decreasing. This could capture negative externalities from search activities with crowding effects, e.g. for profitable arbitrage opportunities in financial markets, or tournaments and races with winner-takes-all compensation in the arts, entertainment, law or R&D. On the other hand, our general framework can also allow for positive externalities, such as spillover effects from entrepreneurial and innovative activities. We revisit the above and other examples in Section 5.

2.3 Income Tax Implementation

We first describe the set of feasible allocations using a Myersonian (1979) direct mechanism and then link this to the implementation through an income tax schedule. In a direct mechanism, individuals announce their type θ and then get assigned observable consumption $c(\theta)$ and total income $y(\theta)$, and unobservable fractions $q^i(\theta) \equiv y_i(\theta)/y(\theta) = r_i(E)\theta_i e_i(\theta)/y(\theta)$ of incomes earned in each activity i . Let $q(\theta) \equiv (q^1(\theta), \dots, q^N(\theta)) \in \Delta^{N-1}$ be the vector of these income shares, where $\Delta^{N-1} \equiv \{q \in \mathbb{R}^N \mid \sum_{i=1}^N q^i = 1, q^i \geq 0\}$.

The incentive constraints that guarantee truth-telling of the agents are:

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

since each type θ can imitate any other type θ' by earning the income of type θ' (and thus getting assigned $c(\theta'), y(\theta')$) using a continuum of effort combinations and hence income shares $p = (p_1, \dots, p_N)$ in the N activities. (Note: we use $e_i = q^i y / (\theta_i r_i(E))$).

The next result is useful for collapsing the incentive constraints (2) into the more standard set of incentive constraints for a screening problem with one-dimensional heterogeneity:

Lemma 1. *In any incentive compatible allocation $\{c(\theta), y(\theta), q(\theta), E\}$, the ratio*

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

with corresponding $\arg \max q(\theta)$, is independent of $(c(\theta), y(\theta))$.

Lemma 1, which generalizes the result for $N = 2$ in Rothschild and Scheuer (2014), 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.

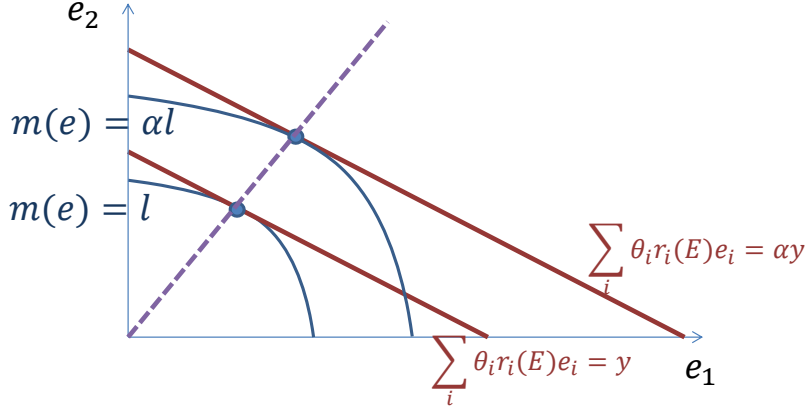


Figure 1: Determination of wages and income shares given E

Moreover, the vector of income shares $q(\theta)$ is chosen so as to minimize the overall effort $m(e)$ subject to achieving a given amount of income: By (3) and linear homogeneity of m ,

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

for any y . By homogeneity and strict quasiconvexity of m , the vector $q(\theta)$ is unique and only depends on E and the vector of skill ratios $\phi \equiv (\theta_1/\theta_N, \dots, \theta_{N-1}/\theta_N) \in \Phi \equiv (0, \infty)^{N-1}$. We therefore write $q_E(\phi)$ (or, with some notational abuse, $q_E(\theta)$) henceforth.⁹

The following lemma, which follows from Berge's Maximum Theorem, states the fact that $q_E(\cdot)$ and $w_E(\cdot)$ are continuous functions of E , which will be useful later.

Lemma 2. $q_E(\theta)$ and $w_E(\theta)$ are continuous in E for all θ .

Figure 1 illustrates the intuition underlying Lemma 1 for the case of two activities. By (4), individuals choose their efforts e_1 and e_2 to minimize their overall effort $m(e)$ subject to achieving a given amount of income y . When the targeted amount of income changes by a factor α , their optimal effort ratio e_1/e_2 remains unchanged, while $m(e)$ increases by the factor α . Hence, income shares q^1 and q^2 and wages $y/m(e)$ are independent of y and only depend on the slope $\theta_1 r_1(E)/\theta_2 r_2(E)$ of the lines in Figure 1.¹⁰

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

⁹Weakly quasiconvex m can be handled with additional notation as in Rothschild and Scheuer (2014).

¹⁰See Section 2.4 for an extension with additional heterogeneity in individuals' preferences for work in the various activities. Note that although homotheticity of m rules out within-individual interactions between earnings and activity-based preferences (though not across individuals), it is perfectly consistent with standard income effects (on the marginal rate of substitution between c and overall income y).

marginal rate of substitution between y and c , $-u_l(c, y/w) / (w u_c(c, y/w))$, is decreasing in w . Then any incentive compatible allocation can be implemented with a non-linear income tax $T(y)$ by the taxation principle (Hammond, 1979, Guesnerie, 1981). As in Rothschild and Scheuer (2013), we can, and henceforth will, restrict attention to allocations $\{c(w), y(w), E\}$ that pool all same-wage individuals at the same (c, y) -bundle.

2.4 Additional Preference Heterogeneity

We briefly point out how the wages $w_E(\theta)$ can be interpreted more broadly. They are defined via the ratio of income to the disutility-of-effort aggregator m . So, as is standard, they can be interpreted either as a literal “dollar per hour” wage or as a measure of the disutility of effort (or as an aggregation of both, as in Choné and Laroque, 2010, and Lockwood and Weinzierl, 2014). Indeed, since the function m is a general aggregator, it may *already* incorporate different psychological or physical effort costs in different activities—for example, costs that might arise from a taste or distaste for certain activities because they are regarded as more or less prestigious, antisocial, etc., or come with other non-pecuniary benefits and burdens. While the preceding formulation, with an individual-independent function m , imposes some uniformity of such tastes across individuals, all our analysis would go through if the function m was individual-dependent, for example, if we had $m(e; \omega)$, where ω captures heterogeneous intrinsic tastes for different types of work.

For concreteness, suppose individuals differ in both their skill vector θ and another unobservable vector ω of intrinsic tastes for each activity. Consider an aggregator $m(e; \omega) = m(e_1/\omega_1, \dots, e_N/\omega_N)$, so ω_i can be interpreted as a measure of the individual’s perception of activity i ’s social “prestige” (if $\omega_i > 1$) or “gaucheness” (if $\omega_i < 1$). Then Lemma 1 goes through with effective wages

$$w_E(\theta, \omega) = \max_{p \in \Delta^{N-1}} m \left(\frac{p_1}{\theta_1 \omega_1 r_1(E)}, \dots, \frac{p_N}{\theta_N \omega_N r_N(E)} \right)^{-1}. \quad (5)$$

For example, suppose $N = 2$ and consider two individuals (θ^a, ω^a) and (θ^b, ω^b) who have the same skills $\theta_1^a = \theta_1^b = \theta_1 < \theta_2 = \theta_2^a = \theta_2^b$ but who differ in how much they value (or notice) the social implications of the two activities, with $\omega_1^a > 1 > \omega_2^a$ and $\omega_1^b = \omega_2^b = 1$. That is, type a perceives activity 1 as prestigious and activity 2 as gauche whereas b has purely pecuniary motives. Then (5) implies that type a will tilt her effort mix towards activity 1 compared to type b even though both have the same skills. Moreover, suppose that $r_2(E)\theta_2 > r_1(E)\theta_1\omega_1^a > r_2(E)\theta_2\omega_2^a$, and that $m(\cdot)$ is sufficiently close to linear so that

types a and b fully specialize in activities 1 and 2, respectively. Then type a 's effective wage is higher than her monetary wage (by factor ω_1^a) but lower than type b 's effective and (equal) dollar wage. In equilibrium, despite the equal skills θ of the two types, type b will earn higher income because of her greater relative preference/tolerance for the highly remunerated activity, while type a will specialize based on her "calling" for the lower paid, prestigious activity. This demonstrates that our model can easily allow for such patterns, which often affect activity choices in practice.

Adding heterogeneity in ω implies pooling an even broader class of individuals at any given effective wage w (and income level). Nevertheless, $w_E(\theta, \omega)$ remains a sufficient statistic for preferences over (c, y) -bundles and all our analysis applies. In particular, all the (distinct) individuals who earn the same income have the same preferences over c and y (except, of course, at bunching points of the tax code).¹¹

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 fact, per Lemma 1, that fixing the vector E determines wages $w_E(\theta)$ and income shares $q_E(\phi)$ makes the problem tractable. Specifically, consider the E -conditional cdf over (w, ϕ) -vectors, given by

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

with the corresponding density $g_E(w, \phi)$.¹² All individuals who earn the same wage w , located on the same iso-wage curve in θ -space as drawn in Figure 2, are pooled in the same allocation (c, y) , and differ only in their relative skills ϕ (so types are fully identified by their (w, ϕ) -combination conditional on E).

We denote the support of the wage distribution for any E by $[\underline{w}_E, \bar{w}_E]$, where $\underline{w}_E \equiv w_E(\underline{\theta}_1, \dots, \underline{\theta}_N)$ and $\bar{w}_E \equiv w_E(\bar{\theta}_1, \dots, \bar{\theta}_N)$. The wage distribution for any given E is simply

¹¹If this were not the case, the techniques developed in Jacquet and Lehmann (2014) could be adapted to incorporate conditional-on-income heterogeneity in our setting.

¹²Extra dimensions of heterogeneity, like the ω discussed in Section 2.4, would add arguments to the cdf, which would subsequently be integrated out.

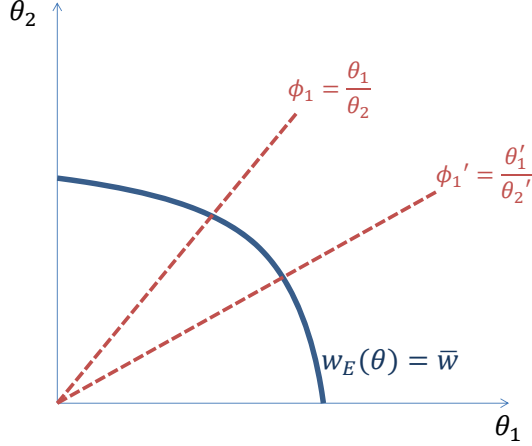


Figure 2: Pooling along iso-wage curves in θ -space conditional on E

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

with the corresponding density $f_E(w) = \int_{\Phi} dG_E(w, \phi)$. We also define the sectoral densities $f_E^i(w) \equiv \int_{\Phi} q_E^i(\phi) dG_E(w, \phi)$; this can be interpreted as an average value of q^i for all wage- w individuals. Clearly, $f_E(w) = \sum_{i=1}^N f_E^i(w)$ for all $w \in [\underline{w}_E, \bar{w}_E]$.¹³ Finally, given any E , we can derive, in an entirely analogous manner, endogenous wage-based Pareto weights over wages $\Psi_E(w)$ and density and sectoral decomposition $\psi_E(w) = \sum_{i=1}^N \psi_E^i(w)$.

The measure F_E^i defined by the cdf $F_E^i(w) \equiv \int_{\underline{w}_E}^w f_E^i(z) dz$ is easily shown to be weakly continuous in E , and analogously for $\Psi_E^i(w) \equiv \int_{\underline{w}_E}^w \psi_E^i(z) dz$.

Lemma 3. *As $E^n \rightarrow E$, $F_{E^n}^i$ converges weakly to F_E^i and $\Psi_{E^n}^i$ converges weakly to Ψ_E^i .*

Finally, allocations $\{c(w), y(w), E\}$ directly imply total effort and utility $l(w) \equiv y(w)/w$ and $V(w) \equiv u(c(w), l(w))$, respectively, as well as the optimal 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 (2013, 2014), we decompose the problem of finding Pareto optimal allocations into two sub-problems. The first 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

¹³In the limiting case with $m(e) = \sum_{i=1}^N e_i$, (3) immediately implies $q_E^i(\phi) \in \{0, 1\}$ almost everywhere, and $w_E(\theta) = \max\{\theta_1 r_1(E), \dots, \theta_N r_N(E)\}$. Then $f_E^i(w)/f_E(w)$ can be interpreted as the share of i -sector workers at w , whereas here it is the i -sector income share at wage w .

allocation for a given E . This inner problem is an almost standard Mirrlees problem; the difference is that the induced vector of aggregate effective efforts has to be consistent with the E that is fixed in the inner problem. For some given Pareto weights $\Psi(\theta)$ (and induced weights $\Psi_E(w)$), we therefore define the inner problem as follows, using $c(V, l)$ to denote the inverse $u(c, l)$ with respect to c :

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

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

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

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

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

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

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

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

It is a well-known result that under single-crossing, the global incentive constraints (11) are equivalent to the local incentive constraints (7) and the monotonicity constraint that income $y(w)$ must be non-decreasing in w .¹⁴

We make two simplifying technical assumptions. First, we abstract from bunching by dropping the monotonicity constraint. Second, we assume that for some sufficiently high \bar{c} and \bar{y} , imposing the additional constraints $c(w) \leq \bar{c}$ and $y(w) \leq \bar{y}$ does not affect the value of problem (6) to (9) for any relevant E . Both assumptions are easily checked *ex post* in computational applications. The former eases the interpretation of the optimal tax formulas we derive below; were it violated, incorporating optimal bunching would

¹⁴See, for instance, Fudenberg and Tirole (1991), Theorems 7.2 and 7.3.

be conceptually no more difficult than in standard Mirrleesian applications. The latter is purely technical: it allows us to easily establish, per the following lemmas, that the supremum in (6) is achieved and that $W(E)$ is upper-semicontinuous in E . It is satisfied, e.g., whenever there is a maximal l , i.e., a \bar{l} such that $\lim_{l \uparrow \bar{l}} u(c, l) = -\infty$.

Lemma 4. *Let $c(V(w), l(w))$ be the unique solution to $u(c, l(w)) = V(w)$, and suppose the constraint set $\{(7), (8), (9), \text{ and } c(V(w), l(w)) \leq \bar{c} \text{ and } l(w) \leq \bar{y}/w\}$ is non-empty. Then some $\{V(w), l(w)\}$ achieves the supremum in (6) over this set.*

Lemma 5. *$W(E)$ is upper semi-continuous at any E^* for which there exists a neighborhood S of E^* and some (\bar{c}, \bar{y}) such that the value of problem (6) to (9) is unaffected on S by the imposition of the additional constraints $c(w) \leq \bar{c}$ and $y(w) \leq \bar{y}$.*

The outer problem is then simply $\sup_E W(E)$. Since $W(E)$ is upper semi-continuous by Lemma 5 (under our technical assumption), the Weierstrass Theorem (viz Luenberger, 1969, p. 40) ensures that this supremum is achieved over any compact subset of $E \in \mathbb{R}^{N+}$. So, if there is a bounded set of feasible E values—for example, if each type's total effort and the returns $r_i(E)$ are bounded—then a solution to the outer problem also exists. In the next two subsections, we characterize the solutions to the inner and outer problems.

3.3 Inner Problem

Solving the inner problem (6) to (9) for a given E yields the following:

Proposition 1. *Given E , the marginal tax rate in any optimum without bunching is such that*

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

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

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

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

$\Psi_E(w) - F_E(w)$. Hence the marginal tax rate is increasing in the degree to which $\Psi_E(w)$ shifts weight to low-wage individuals compared to $F_E(w)$.

The only difference from standard formulas is that, at each wage, the marginal keep shares $1 - T'(y(w))$ are adjusted by the factor $1 - \sum_{i=1}^N (f_E^i(w)/f_E(w))(\zeta_i/r_i(E))$. As we will argue in the next Section 3.4, this factor is a local correction for the general equilibrium effects and externalities caused by income earned by wage w -individuals. In particular, the multiplier ζ_i on the i^{th} constraint (8) is the optimal correction on effective effort in sector i —i.e., the correction taking general equilibrium effects into account. The term $\sum_{i=1}^N (f_E^i(w)/f_E(w))(\zeta_i/r_i(E))$ is therefore an income-share weighted average of the general equilibrium corrections ζ_i/r_i on the incomes earned in the various activities.

3.4 Outer Problem

In this section, we characterize the optimal corrections ζ_i using the conditions for an optimal E from the outer problem. In particular, we are interested in the relationship between the general equilibrium corrections ζ_i/r_i and the partial equilibrium Pigouvian taxes τ_p^i that would align the social and private marginal products of income earned in activity i , defined by $r_i(E)(1 - \tau_p^i(E)) \equiv \partial Y(E)/\partial E_i$.¹⁵

We derive necessary conditions for the outer problem using a Lagrange formulation.¹⁶ By the envelope theorem, changes in E have a direct welfare effect through the left-hand-side of the consistency constraints (8) and through their effects on the (sectoral) wage distributions in the resource constraint (9) (consistency constraints (8)) and objective (6). These distributional effects are quite complex, and it is more fruitful to characterize (and then integrate) the individual-level welfare effects of a change in E , effects which arise because individuals' wages and across-activity effort compositions change as the returns $r_i(E)$ change.

Formally, we divide the marginal welfare effects of a small change dE_i in E_i into four classes: (i) the direct effect on the left-hand side of the i^{th} consistency constraint (8) and three other effects which capture the effect on any given type θ . The change dE_i changes type θ 's wage. We designate by (ii) the direct effects that this wage change has on (6) to (9), holding fixed the type's effort $l(\theta)$ and utility $V(\theta)$. We designate by (iii) the indirect effects that this wage change has on θ 's $(l(\theta), V(\theta))$ -allocation as her wage change induces her to move along the fixed schedules $(l(w), V(w))$. Finally, since dE_i changes

¹⁵ τ_p^i would be the optimal correction on activity- i income if activity-specific instruments were available (see Rothschild and Scheuer, 2014); thus, τ_p^i is the standard Pigouvian tax under perfect targeting.

¹⁶This requires—and we assume—that W' exists and coincides with the derivative of the associated Lagrangian. Appendix B provides simple sufficient conditions under which this is generically true.

the returns $r_i(E)$ to effort in the N activities, type θ 's optimal across-activity allocation of efforts $e_i(\theta)$ will change for any given total effort $l(w)$. We designate by (iv) the welfare effects (through the consistency constraint (8)) of this re-allocation.

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

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

Using the envelope theorem and denoting the semi-elasticity of $r_j(E)$ w.r.t. E_i by

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

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

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

i.e., the income-share weighted average of the return semi-elasticities. For an individual with original wage w and original income share vector q_E , the change in the wage induced by a change dE_i in E therefore causes a change $l'(w)w \sum_{j=1}^N q^j \beta_i^j(E) dE_i$ in l . The average change in l for all types with original wage w is therefore

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

where $\mathbb{E}[q_E^j(\phi) \mid w] = \int_{\Phi} q_E^j(\phi) g_E(\phi \mid w) d\phi$ is the average of q^j over the set $\{\theta \mid w_E(\theta) = w\}$ of all wage- w individuals. We can “undo” this w -specific change in average effort l by modifying the l schedule to $\tilde{l}(w) = l(w) - l'(w)w \delta_E^i(w) dE_i$, where

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

Analogously, we can modify the V -schedule to $\tilde{V}(w) = V(w) - V'(w)w \delta_E^i(w) dE_i$ in order to “undo” the w -specific change in average welfare V . Performing these modifications jointly with dE_i greatly simplifies the outer problem effects (iii) by removing any average effort and utility effects for the set of types at each wage w . In fact, these modifications

also ensure that average consumption is unchanged at each w .¹⁷ By the envelope theorem, these schedule modifications have no welfare effects at the margin.

3.4.1 Redistributive Effects

The objective (6) changes because individuals' changing wages move them along the $V(w)$ schedule (i.e., from effect (iii)). By analogy to (16), the effect of dE_i for a given schedule is simply $V'(w)w \sum_{j=1}^N \psi_E^j(w) \beta_i^j(E) dE_i$. Adding this to the welfare effect of the change in the V -schedule to \tilde{V} , namely $-V'(w)w \delta_E^i(w) dE_i$, yields

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

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

Note that $\sum_{j=1}^N R_j(E) = 0$; intuitively: the R_j capture w -specific re-allocations of utility across workers with different sectoral intensities q . For the same reason, each R_j disappears in the natural benchmark with equal welfare weight on all individuals with 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 result e.g. from relative Pareto weights $\Psi(\theta) = \tilde{\Psi}(F(\theta))$).¹⁸ Otherwise, if, e.g., dE_i increases the relative returns to activities in which workers with a high relative welfare weight earn much of their income, the resulting re-allocation in utilities is welfare enhancing.

3.4.2 Incentive Constraint Effects

There are no incentive effects of dE_i given the fixed schedules l and V (individuals move along an incentive compatible schedule). The schedule modification to (\tilde{l}, \tilde{V}) is readily shown to change $V'(w) - u_l(c(w), y(w))l(w)/w$ by $-V'(w)w \delta_E^i(w) dE_i$.¹⁹

Using (17) and integrating over all wages, the incentive effects from (iii) are therefore

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

¹⁷To wit, dropping the common argument w and using (7) and (55) yields $\tilde{c} - c = c(\tilde{V}, \tilde{l}) - c(V, l) = \frac{1}{u_c}(\tilde{V} - V) - \frac{u_l}{u_c}(\tilde{l} - l) = \left(\frac{1}{u_c} V' - \frac{u_l}{u_c} l' \right) w \delta_E^i = \left(-\frac{u_l}{w u_c} + \frac{u_c c' + u_l l'/w}{u_c} \right) w \delta_E^i = c' w \delta_E^i$.

¹⁸In this benchmark, the planner only cares about wage inequality without inherent sectoral preferences.

¹⁹Use $\tilde{V}'(w) = V'(w) - V'(w)w \frac{d\delta_E^i(w)}{dw} dE_i - \frac{d(V'(w)w)}{dw} \delta_E^i dE_i$ and

$$\frac{1}{w} (u_l(\tilde{c}(w), \tilde{l}(w))\tilde{l}(w) - u_l(c(w), l(w))l(w)) = \frac{d(u_l(c(w), l(w))l(w))}{dw} \delta_E^i dE_i = \frac{d(V'(w)w)}{dw} \delta_E^i dE_i.$$

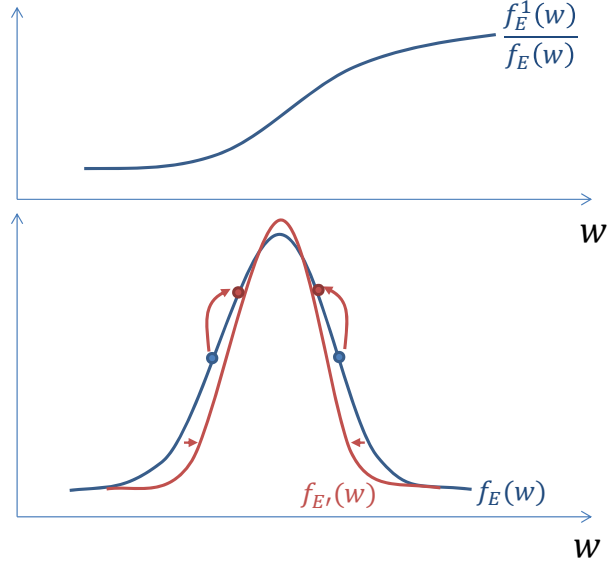


Figure 3: Incentive constraint effects

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

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

As with the R_j , $\sum_{j=1}^N I_j(E) = 0$. To interpret the terms I_j , suppose $\eta(w) > 0$ (i.e., down-binding incentive constraints). Then $dE_i > 0$ is welfare reducing (respectively, increasing) if it increases (decreases) the returns to activities j with $d \left(\frac{f_E^j(w)}{f_E(w)} \right) / dw > 0$, i.e., to activities that are locally associated with high wages. This is because dE_i makes the wage distribution more (less) unequal in this case, which tightens (loosens) the local incentive constraints. The effect is therefore a generalized version of the one pointed out by Stiglitz (1982) for a two-type model with two sectors and fixed activity choice.

Figure 3 illustrates this for $N = 2$, so (20) becomes $-\lambda(\beta_i^1 - \beta_i^2)I_1$. If sector 1 is the high-wage sector and $\beta_i^1 - \beta_i^2 < 0$ (i.e., relative returns in activity 1 fall with E_i), then $dE_i > 0$ compresses the wage distribution. This is because higher wage individuals—who on average have a greater effort intensity in activity 1—see their wages fall, on average, relative to lower wage, activity 2-intensive individuals. This local wage compression is welfare improving if downward-redistribution is desirable (i.e., $\eta(w) > 0$). The total effect in (20) is the integral over all these local wage compression effects.

3.4.3 Resource Constraint Effects

A wage change induced by dE_i affects (9) in two ways. First, given $l(w)$ and $c(w)$, the change in w *directly* affects (9) via the w appearing in the integrand (effect (ii) above). Second, a change in w moves an individual along the $l(w)$ and $c(w)$ schedules. This second effect is, by construction, exactly cancelled on average at each wage w by the schedule variation to (\tilde{l}, \tilde{V}) . Using (17), the overall effect from the first change is simply

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

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

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

(If, e.g., activity i effort raises the returns to the various activities, it generates positive externalities and the Pigouvian tax is negative.) Using (23) in (22) yields a resource constraint effect of simply $-\lambda t_p^i(E) dE_i$. That is: dE_i increases (decreases) welfare through the resource constraint if and only if it generates positive (negative) externalities.

3.4.4 Consistency Constraint Effects

Next, consider the effects of dE_i on consistency constraint j . First, there is the direct effect (i.e., (i)), which is $\lambda \tilde{\zeta}_j$ if $i = j$ and 0 otherwise. Hence the sum over all constraints is simply $\lambda \tilde{\zeta}_i$. Second, there are various effects on the right-hand side of constraint j . These can be written, compactly, as a sum of two terms. The first is

$$-\lambda \tilde{\zeta}_j \sum_{k=1}^N \beta_i^k(E) C_{kj}(E), \quad (24)$$

where
$$C_{kj}(E) \equiv \frac{1}{r_j(E)} \int_{\underline{w}_E}^{\overline{w}_E} w^2 l'(w) \text{Cov} \left(q_{E'}^j, q_E^k \mid w \right) f_E(w) dw \quad (25)$$

with $\text{Cov} (q^j, q^k) \equiv \mathbb{E} [q^j q^k] - \mathbb{E} [q^j] \mathbb{E} [q^k]$. The second is

$$-\lambda \tilde{\zeta}_j \sum_k \beta_i^k(E) S_{kj}(E) \quad (26)$$

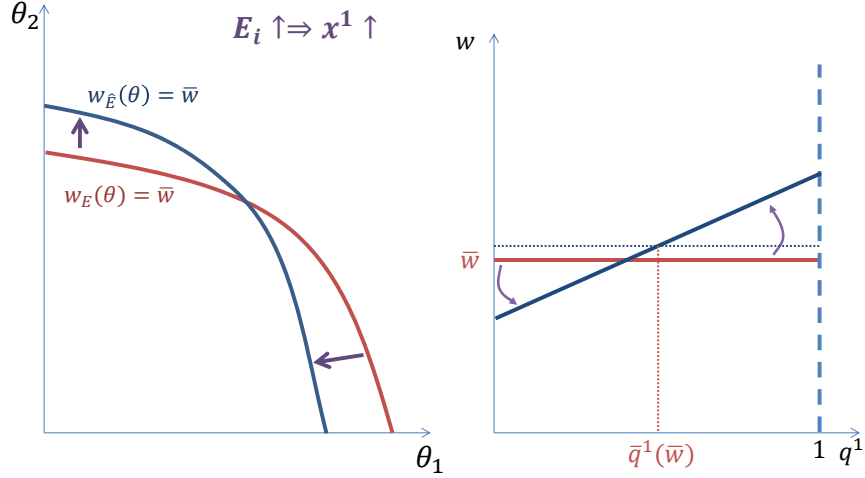


Figure 4: Overall effort re-allocation effect

with

$$S_{kj}(E) \equiv \frac{1}{r_j(E)} \int_{\underline{w}_E}^{\bar{w}_E} y(w) \int_{\Phi} Q_k^j(x_E(\phi)) x_E^k(\phi) dG_E(w, \phi), \quad (27)$$

where $Q_k^j(x_E(\phi))$, defined formally below, measures the change in a given type's sector- j income share that is induced by that type's optimal adjustment of her sectoral effort ratios in response to an increase in the returns to sector k .

For readers who wish to skip the formal derivation of these terms, which follows below, we first provide some intuition. The term (25) is an *effort reallocation effect*. The intuition is tightly linked to our chosen schedule change from $l(w)$ to $\tilde{l}(w)$. This zeroes out the *average* change in aggregate effort l at any given w , but, of course, some individuals originally pooled at wage w will see their wage, and hence their l , rise (if $l'(w) > 0$), while others will see it fall. This re-allocates effort across individuals and therefore, since different individuals at w have different effort intensities in the various activities, across activities. If the activity- j income share q^j is uncorrelated with this effort change at any given w , then activity- j effort will also remain unchanged. If it is positively correlated, however, then activity- j effort increases, and vice versa. In particular, if $dE_i > 0$ increases the returns to activities k , and if individuals who have a high income share in k also have a high income share in j , then individuals with a high q^j on average see their wage increase more than proportionally. If $l'(w) \geq 0$, this effectively shifts effort towards activity j and increases the RHS of consistency constraint j . Expression (24) will reflect this via $C_{kj} > 0$.

Figure 4 illustrates the effort reallocation effect for $N = 2$. Suppose $dE_i > 0$ increases the relative return to activity 1, as drawn in the figure, so $\beta_i^1 - \beta_i^2 > 0$. Individuals on the iso-wage curve $w_E(\theta) = \bar{w}$ and with a high ratio θ_1/θ_2 (and thus a high intensity q^1 in activity 1) will experience a rise in their wage relative to those with a low θ_1/θ_2 . If

$l'(w) \geq 0$, then $C_{11}, C_{22} \geq 0$, and the high (low) θ_1/θ_2 types will increase (decrease) their efforts l . Effort will thus effectively flow out of sector 2 and into sector 1. Of course, if every wage- \bar{w} earner had the same q^1 (so $\text{Var}(q^1|\bar{w}) = 0$), then the rotation shown in the right panel would be without consequences and C_{11} and C_{22} would be zero.

The second term, (26), captures an *activity shift effect*. Intuitively, suppose $dE_i > 0$ increases the returns r_k to an activity k (so $\beta_i^k(E) > 0$). Such a change will lead each individual to re-optimize her relative efforts across the various activities—certainly towards k , but potentially also across other activities. By definition, $Q_k^j > 0$ in those activities j which experience a relative increase in effort. If $Q_k^j > 0$ on average across all individuals, then $S_{kj} > 0$, and the effect of dE_i on r_k causes a net shift of effort into activity j and increases the RHS of the j -th consistency constraint. This is illustrated in Figure 5 for $N = 2$: If $dE_i > 0$ increases the relative return r_1/r_2 to activity 1, individuals will shift their effort allocation e_1/e_2 towards activity 1.

We now derive these two effects formally (the casual reader can skip directly to Subsection 3.4.5). For this derivation, it is useful to rewrite consistency constraint j following (10) as $E_j = \int_{\Theta} \theta_j e_j(\theta) dF(\theta)$ and to note that

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

by homogeneity of degree 1 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*

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

I.e., $\zeta \equiv (\zeta_1, \dots, \zeta_{N-1}, 1)$ is homogenous of degree 0 in the return vector $(\theta_1 r_1(E), \dots, \theta_N r_N(E))$.

The effective effort integrated over on the RHS of consistency constraint j is

$$\theta_j e_j(\theta) = l(w) \theta_j \Omega_j(\zeta(x_E(\phi))) \quad \text{with} \quad \Omega_j(\zeta(x_E(\phi))) \equiv \frac{\zeta_j(x_E(\phi))}{m(\zeta(x_E(\phi)))}. \quad (28)$$

The left-hand equation motivates a decomposition of the effect into the change in the overall level of effort $l(\cdot)$ for each individual (part of (iii)), holding constant the cross-sectoral allocation of efforts Ω , and second, changes in Ω , which reflect a re-allocation of effort across sectors due to the change in the relative returns x_E (effect (iv)). The former will yield (24) and the latter (26).

Overall effort re-allocation effect. The direct effect on $l(\cdot)$ for type ϕ individuals with

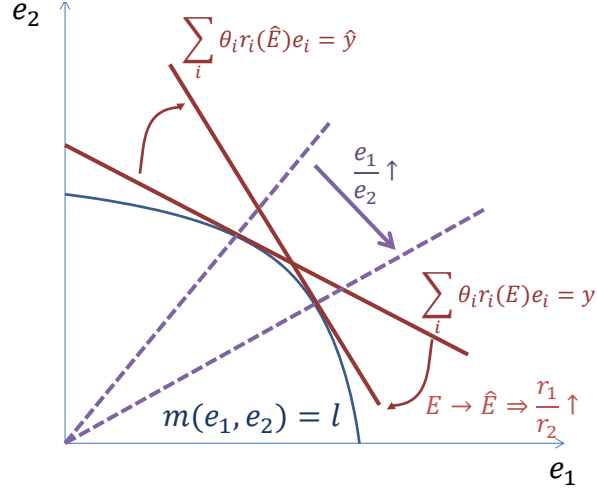


Figure 5: Activity shift effect

wage w is $wl'(w) \sum_{k=1}^N q_E^k(\phi) \beta_i^k(E)$, and, using $\theta_j \Omega_j = \theta_j e_j / l = wq^j / r_j$, the effect on (28) is

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

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

$$\frac{1}{r_j(E)} \sum_{k=1}^N \beta_i^k(E) w^2 l'(w) \mathbb{E} \left[q_E^j(\phi) q_E^k(\phi) \mid w \right]. \quad (29)$$

Changing from l to \tilde{l} changes the average $l(\cdot)$ at w by $-wl'(w) \sum_{k=1}^N \mathbb{E} [q_E^k(\phi) \mid w] \beta_i^k(E)$, and so the average change in sector j equivalent effort in (28) is

$$-\frac{1}{r_j(E)} \sum_{k=1}^N \beta_i^k(E) w^2 l'(w) \mathbb{E} \left[q_E^j(\phi) \mid w \right] \mathbb{E} \left[q_E^k(\phi) \mid w \right]. \quad (30)$$

Adding (29) and (30) and integrating over all wages yields (24).

Activity shift effect. The effect of dE_i through the change in the vector of effort ratios ζ on (28) is,²⁰

$$l(w) \theta_j \sum_{k=1}^N \sum_{l=1}^N \frac{\partial \Omega_j(\zeta(x_E(\phi)))}{\partial \zeta_l} \frac{\partial \zeta_l(x_E(\phi))}{\partial (\theta_k r_k(E))} \frac{\partial \theta_k r_k(E)}{\partial E_i}. \quad (31)$$

²⁰Our assumptions on m ensure that $\zeta_l(x_E(\phi))$ is continuous and differentiable almost everywhere—with potential non-differentiabilities only at corners where $\zeta_l(x_E(\phi)) = 0$.

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

$$Z_j(x_E(\phi)) \equiv \frac{r_j(E) \theta_j}{w} = x_E^j(\phi) \min_{p \in \Delta^{N-1}} m \left(\frac{p_1}{x_E^1(\phi)}, \dots, \frac{p_{N-1}}{x_E^{N-1}(\phi)}, p_N \right), \quad (32)$$

where we used (3) and homogeneity of degree 1 of m . Define

$$Q_k^j(x_E(\phi)) \equiv Z_j(x_E(\phi)) \sum_{l=1}^N \frac{\partial \Omega_j(\zeta(x_E(\phi)))}{\partial \zeta_l} \frac{\partial \zeta_l(x_E(\phi))}{\partial (r_k(E) \theta_k)} \theta_N r_N(E) \quad (33)$$

and substitute into (31) to yield:

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

Integrating over all wages and all ϕ yields (26).

It is worth noting briefly that Q_k^j is *not* the *total* change in q^j induced by a change in the returns to k . Indeed, $q_E^j(\phi) = Z_j(x_E(\phi)) \Omega_j(\zeta(x_E(\phi)))$, so a change in the returns to k can be thought of as having two effects on q^j : a *mechanical* effect through the changes in returns Z_j , and an activity shift effect through the change in Ω_j . Only the latter effect appears in expression (33) for Q_k^j and in the welfare effect formula (26).

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

Lemma 6. (i) $\sum_{j=1}^N r_j(E) C_{kj}(E) = \sum_{j=1}^N r_j(E) S_{kj}(E) = 0$ for all $k = 1, \dots, N$.
(ii) $\sum_{k=1}^N C_{kj}(E) = \sum_{k=1}^N S_{kj}(E) = 0$ for all $j = 1, \dots, N$.

The intuition for part (i) of Lemma 6 hinges on the fact that the $r_j(E)$ -weighted sum of the right-hand sides of the N constraints in (8) is $\int_{\bar{w}_E}^w y(w) f_E(w) dw$, and the sectoral composition of income at wage w is irrelevant for this sum. Indeed, both effects reflect income shifts across activities in response to return changes and thus have to add up to zero. The adding-up property in part (ii) of the lemma is a consequence of the fact that proportional changes in *all* returns (a) do not affect individuals' cross-sectoral effort allocation since m is linear homogeneous, so there are no activity shift effects, and (b) cause equi-proportional changes in the wages of all types θ , and hence no cross-sectoral re-allocation of overall effort $l(w)$ at any wage.

As a direct consequence of Lemma 6, (24) and (26) are non-zero only if a change in E_i affects *relative* returns.

3.4.5 Putting All Together

To find the total welfare effect of a marginal change in E_i , we combine (18), (20), (22), with (24), (26) and the direct effect ξ_i . Moreover, because of the adding-up property in Lemma 6 (ii) and the fact that $\sum_j I_j = \sum_j R_j = 0$, only *relative* return changes matter for the effects (18), (20), (24) and (26), so we can equivalently write (18) as $-\lambda \sum_j (\beta_i^j - \beta_i^N) R_j$ and analogously for the others. Defining $\Delta\beta_i^j(E) \equiv \beta_i^j(E) - \beta_i^N(E) = \left(x_E^j(\phi)\right)^{-1} \partial x_E^j(\phi) / \partial E_i$ (i.e., the relative return semi-elasticity), we summarize the results from this subsection:

Lemma 7. *At any $E_i > 0$, the welfare effect of a marginal change in E_i is*

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

with $R_j(E)$, $I_j(E)$, $t_p^i(E)$, $C_{jk}(E)$ and $S_{jk}(E)$ respectively given by (19), (21), (23), (25) and (27).

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 .

3.5 Marginal Tax Rate Results and Outer Problem Dimensionality

We are now ready to characterize the optimal corrections ξ_i in the marginal tax rate formula (12) and compare them to the Pigouvian tax rates t_p^i . Using Lemma 7, the N interior optimality conditions $\partial W / \partial E_i = 0$ can be written compactly as:

$$(\mathbf{I}_N - \Delta\beta(\mathbf{C} + \mathbf{S})) \vec{\xi} = \vec{t}_p + \Delta\beta (\vec{I} + \vec{R}), \quad (34)$$

where \mathbf{I}_N denotes the $N \times N$ identity matrix, $\Delta\beta$, \mathbf{C} , and \mathbf{S} are the matrices with $(i, j)^{\text{th}}$ elements $\Delta\beta_i^j(E)$, $C_{ij}(E)$, and $S_{ij}(E)$, respectively, and \vec{I} , \vec{R} , $\vec{\xi}$ and \vec{t}_p are the column vectors with elements $I_i(E)$, $R_i(E)$, ξ_i , and $t_p^i(E)$, respectively.

Let \vec{n} denote the column vector with i^{th} element $n_i = (1/r_i(E))(f_E^i(w)/f_E(w))$ and \vec{n}' its transpose. Fix any wage w and consider a small change in the tax code at $y(w)$ such that wage- w individuals are induced to increase their earnings by a small amount dy . Ignoring general equilibrium effects, $dY_i = (f_E^i(w)/f_E(w)) dy$ is the induced change

in sector- i income and $dE_i = (1/r_i(E)) (f_E^i(w)/f_E(w)) dy = n_i dy$ the induced change in aggregate activity- i effort. The vector \vec{n} thus denotes the (partial-equilibrium) directional change in E that would be induced by a small variation in the tax code at $y(w)$. The term $\vec{n}'\Delta\beta$ denotes the relative return changes induced by such a variation.

If $\vec{n}'\Delta\beta = 0$, so this variation has no relative return effects, then left-multiplying (34) yields $\vec{n}'\vec{\zeta} = \vec{n}'\vec{t}_p$, i.e.,

$$\sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \frac{\zeta_i}{r_i(E)} = \sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \tau_p^i(E). \quad (35)$$

The RHS of (35)—the income share-weighted average of the Pigouvian taxes on incomes in the N activities—is the partial-equilibrium corrective tax. The LHS is the optimal correction in the income tax formula (12), i.e., the optimal *general equilibrium* correction. We conclude that the general and partial equilibrium corrections coincide precisely at income levels at which small changes in the marginal tax rate would induce no relative return effects. When there *are* relative return effects, and $\vec{n}'\Delta\beta \neq 0$, then the optimal correction, per (12) and (34), will generically diverge from the partial equilibrium correction.

The following result provides a simple characterization of when the partial and general equilibrium corrections coincide, and are both equal to zero, so the marginal tax rate formula (12) is the same as in a standard Mirrlees model:

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

(Here, β denotes the matrix with elements $\beta_i^j(E)$.) Let $N - K$ denote the rank of mapping $E \rightarrow r(E) = (r_1(E), \dots, r_N(E))'$ and hence of β . Since the return vector $r(E)$ is a sufficient statistic for individual behavior, conditional on a given tax code, 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.

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

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

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 magnitude (and possibly sign). Then there is a two-dimensional plane with directions of no relative return effects and no externalities spanned by the vectors $(-a, 1, 0)$ and $(-b, 0, 1)$. The vector orthogonal to both is $(1, a, b)$, so $\tilde{E}_1 = E_1 + aE_2 + bE_3$ is a sufficient statistic for the return vector $r(E)$. The outer problem can again be written with a single consistency constraint, namely a weighted average of the three consistency constraints in (8).

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

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

Finally, for the system of optimality conditions (34) to uniquely identify the vector $\vec{\zeta}$, the matrix $\mathbf{A} \equiv \mathbf{I}_N - \Delta\beta(\mathbf{C} + \mathbf{S})$ needs to be non-singular at the optimum. We assume this in the following.

4 Two Sectors

If $N = 2$, the adding up properties of Lemma 6 can be used to solve the system of optimality conditions (34) for $\vec{\zeta}$ explicitly.

Lemma 8. *At any Pareto optimum with $N = 2$,*

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

where

$$C(E) \equiv \int_{\underline{w}_E}^{\bar{w}_E} w^2 l'(w) \text{Var}(q_E^1|w) f_E(w) dw, \quad (37)$$

$$S(E) \equiv \int_{\underline{w}_E}^{\bar{w}_E} y(w) \int_{\Phi} Q_1^1(x_E^1(\phi)) x_E^1(\phi) dG_E(w, \phi), \quad (38)$$

and

$$\gamma_2(E) = 1 + \left(\frac{\Delta\beta_2^1(E)}{r_2(E)} - \frac{\Delta\beta_1^1(E)}{r_1(E)} \right) (C(E) + S(E)), \quad (39)$$

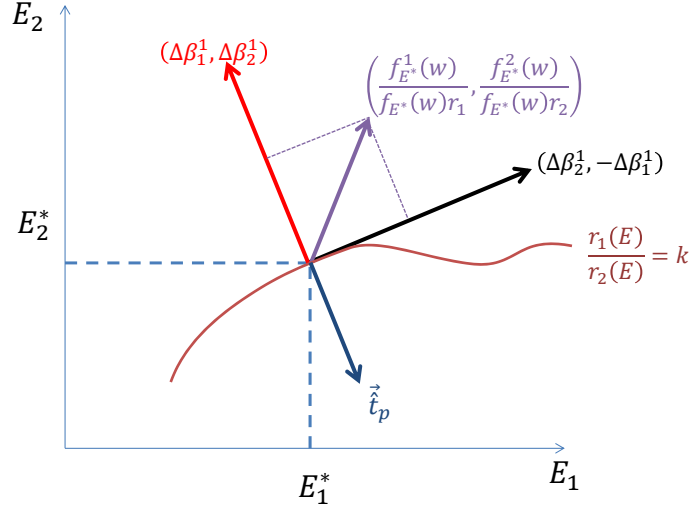


Figure 6: Adjustment factor and directions of no and maximal relative return effects

The system (36) 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 3.5, 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 (36) and the angle between \vec{n} and the direction of no relative return effects $(\Delta\beta_2^1(E), -\Delta\beta_1^1(E))$.²¹ This is illustrated in Figure 6, which shows an iso-relative return curve in (E_1, E_2) -space as well as the (tangent) direction of no relative return effects, the (perpendicular) direction of maximal relative return effects, and the projection of the vector \vec{n} on the latter, all starting from a Pareto optimum (E_1^*, E_2^*) .

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: $\vec{t}_p = \chi(\Delta\beta_1^1, \Delta\beta_2^1)$ for some χ , as depicted in Figure 6. Clearly, this is trivially the case when there are no externalities (so that $\chi = 0$, as in Rothschild and Scheuer, 2013) or only one activity affects returns (so that $\Delta\beta_2^1 = t_p^2 = 0$, see Rothschild and Scheuer, 2014), as we will discuss in more detail below.

²¹The denominator $\gamma_2(E)$ is the eigenvalue associated with the eigenvector $(\Delta\beta_1^1(E), \Delta\beta_2^1(E))'$ of A (which is also the direction of *maximal* relative return effects), so it is non-zero by our assumption that A is non-singular. Moreover, $\gamma_2(E) > 0$ if E is stable in the sense of Appendix C.

5 Applications

In this section, we illustrate how our general framework can provide useful insights into the shape of the optimal income tax schedule in a number of important applications. These include standard constant returns to scale economies (CRS) with multiple sectors, economies with increasing or decreasing returns to scale, CRS economies where returns deviate from marginal products, and non-CRS economies where one of the N activities generates or bears externalities, as discussed in Section 2.2.

5.1 No Externalities

We begin with the externality-free case where $Y(E)$ has CRS and $r_i(E) = \partial Y(E)/\partial E_i$ for all i , so social and private returns coincide. Rothschild and Scheuer (2013) consider the special case with $N = 2$. The tools from Section 3 can be used to investigate the novel effects that arise when activity choice is along more than one margin. The simplest way to shed light on this is to add a third, linear sector, leading to the production function $Y(E) = \hat{Y}(E_1, E_2) + E_3$, where \hat{Y} has CRS and $\partial^2 \hat{Y}/\partial E_i^2 < 0$, $i = 1, 2$. This is particularly tractable as E_3 has no effects on any returns and $\rho \equiv E_1/E_2$ remains a sufficient statistic for the wage distribution (as $r_i = \partial Y(E)/\partial E_i$). The general system of optimality conditions (34) can be solved to obtain the following modified marginal tax rate adjustment factor.

Proposition 4. *If $N = 3$, $Y(E) = \hat{Y}(E_1, E_2) + E_3$, \hat{Y} is homogenous of degree 1 and $r_i(E) = \partial Y(E)/\partial E_i$, $i = 1, 2, 3$, then the numerator in the marginal tax rate formula (12) is*

$$1 - \sum_{i=1}^3 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i(E)} = 1 + \frac{f_E^1(w) + f_E^2(w)}{f_E(w)} \left(\frac{f_E^1(w)}{f_E^1(w) + f_E^2(w)} - \hat{\alpha}(\rho) \right) \hat{\xi} \quad (40)$$

with

$$\hat{\xi} = -\frac{1}{r_1(1 - \hat{\alpha})} \frac{\beta_1^1(I_1 + R_1) + \beta_1^2(I_2 + R_2)}{1 - \beta_1^1(C_{11} + S_{11} - \rho(C_{12} + S_{12})) - \beta_1^2(C_{21} + S_{21} - \rho(C_{22} + S_{22}))}, \quad (41)$$

where $\hat{\alpha}(\rho) \equiv Y_1/\hat{Y}$ is the activity 1 share of the combined incomes of activities 1 and 2.

In the two-sector case where E_3 and $f_E^3(w)$ vanish, Lemma 6 implies $C_{21} = -C_{11}$, $r_1 C_{12} = -r_2 C_{11}$, and $r_2 C_{22} = r_1 C_{11}$, and analogously for the S_{ij} terms. Similarly, $I_2 + R_2 = -(I_1 + R_1)$. Denoting by $\sigma(\rho)$ the substitution elasticity of $Y(E)$ and by $\alpha(\rho) \equiv Y_1(E)/Y(E)$ the aggregate income share of sector 1, and using the definitions of C and S in (37) and (38), the adjustment factor then collapses to the formula from Rothschild and Scheuer (2013):

Corollary 1. *If $N = 2$, $Y(E)$ is homogenous of degree 1 and $r_i(E) = \partial Y(E)/\partial E_i$ then the numerator in the marginal tax rate formula (12) is*

$$1 - \sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\zeta_i}{r_i(E)} = 1 + \left(\frac{f_E^1(w)}{f_E(w)} - \alpha(\rho) \right) \bar{\zeta} \quad \text{with} \quad \bar{\zeta} \equiv \frac{(I_1 + R_1) / \sigma}{\alpha(1 - \alpha)Y + (C + S) / \sigma}. \quad (42)$$

This corollary implies a regressive adjustment to standard Mirrleesian tax rates. Intuitively, lower taxes at wages where the high-wage activity (say activity 1) is prevalent will encourage effort there. By complementarity, this increased effort increases the relative returns to the low-wage activity, which is desirable under typical social preferences. This effect is reflected in the re-distributional terms I_1 and R_1 in the numerator of $\bar{\zeta}$. The increase in relative returns to the low-wage activity is partially counteracted, however, since it induces individuals to shift effort out of the high-return activity into the lower-return activity. This is captured by the reallocation effects C and S , which blunt the regressive adjustment. The optimal tax schedule is therefore more regressive than in a Mirrlees (1971) model with fixed wages, but less regressive than in an endogenous wage model with fixed occupations, such as Stiglitz (1982).

The adjustment disappears at wage levels w where $f_E^1(w)/f_E(w) = \alpha$, so that the local and aggregate income shares coincide. This reflects the discussion in Section 3.5: at such points, $\vec{n}' = (f_E^1(w)/(f_E(w)Y_1), f_E^2(w)/(f_E(w)Y_2))$ reduces to $(E_1/Y, E_2/Y)$ and therefore points in the direction $(\rho, 1)$ in which there are zero relative return effects which, here, is trivially also a direction of zero externalities.

When the third sector is active, the adjustment factor in (40) similarly vanishes whenever $f_E^1(w)/(f_E^1(w) + f_E^2(w)) = \hat{\alpha}$. The factor is now scaled down by the local share $(f_E^1(w) + f_E^2(w))/f_E(w)$ of income earned in sectors 1 and 2, however, reflecting the fact that the relative return effects only operate through a fraction of the population. Moreover, the term $\hat{\zeta}$ in (41) will generally diverge from the term $\bar{\zeta}$ in (42). This is because, although the formula for $\hat{\zeta}$ in (41) is the same in a two- and a three-sector model, the adding up properties in Lemma 6 that pin down the relationship between the four $C_{ij} + S_{ij}$, $i, j = 1, 2$ terms in a two-sector model are less informative with a third sector. Nevertheless, it is instructive to use Lemma 6 to re-write (41) as

$$\hat{\zeta} \equiv \frac{(I_1 + R_1) / \hat{\sigma} + E_1 \beta_1^2 (I_3 + R_3)}{\hat{\alpha}(1 - \hat{\alpha})\hat{Y} + \frac{\hat{C}_{11} + \hat{S}_{11}}{\hat{\sigma}} + \hat{\alpha} \frac{\hat{C}_{13} + \hat{S}_{13}}{\hat{\sigma}} + \hat{\alpha}(1 - \hat{\alpha})E_1 \beta_1^2 \left[\frac{\hat{C}_{31} + \hat{S}_{31}}{\hat{\alpha}} - \frac{\hat{C}_{32} + \hat{S}_{32}}{1 - \hat{\alpha}} \right]}, \quad (43)$$

where $\hat{\sigma}(\rho)$ is the substitution elasticity of \hat{Y} and $\hat{C}_{ij} \equiv r_j C_{ij}$, $\hat{S}_{ij} \equiv r_j S_{ij}$. Comparing with (42) reveals an extra term in the numerator and two in the denominator.

First, consider the extra term $\hat{\alpha}(\hat{C}_{13} + \hat{S}_{13})/\hat{\sigma}$ in the denominator. With two sectors, any outflow $-(\hat{C}_{11} + \hat{S}_{11})$ of sector 1 earnings (caused by the r_1 -decrease associated with increased E_1) is necessarily an inflow into sector 2; with a third sector, some of the outflow will instead go to sector 3. Unlike flows into sector 2, however, sector 3 inflows do not decrease $\rho = E_1/E_2$, and therefore do not blunt the indirect redistribution achieved by an increase in E_1 . Relative to the two-sector model, the presence of the third sector, reflected in this term, makes $\hat{\zeta}$ larger and the optimal tax more regressive.

Next, consider the second extra term in the denominator. With two activities, a proportional increase in both r_1 and r_2 induces no activity shifts and hence has no effect on ρ . With three activities, however, a proportional increase in r_1 and r_2 induces an income shift out of activity 3 and into activities 1 and 2, which affects ρ insofar as a change in r_3 leads to unequal percentage changes in sector 1 and 2 incomes (which, by Lemma 6 (ii), implies the same for a proportional and simultaneous change in r_1 and r_2). This is reflected in the factor in square brackets in (43) (and the fact that $\beta_1^2 > 0$). For example, if $(\hat{C}_{31} + \hat{S}_{31})/\hat{\alpha} < (\hat{C}_{32} + \hat{S}_{32})/(1 - \hat{\alpha}) < 0$, activity 1 income changes by relatively more than activity 2 income when r_1 and r_2 change proportionally. This means that the effects of r_1 outweigh the effects of r_2 , which again reinforces the regressive adjustment from the preceding paragraph.

Finally, towards understanding the extra term in the numerator, ignore R_3 and suppose that sector 1 is the high income sector, so $f_E^1(w)/f_E(w)$ is increasing in w . With two activities, this would mechanically imply a decreasing $f_E^2(w)/f_E(w)$. The wage changes induced by an increase in E_1 would then beneficially redistribute, by reducing r_1 and increasing r_2 , from the high income activity 1 to the low income activity 2. If there is a third activity with increasing $f_E^3(w)/f_E(w)$ (and hence $I_3 > 0$), however, then sector 2 is *even more* of a low income sector, and the indirect redistributive benefits of an increase in E_1 are magnified—and so is $\hat{\zeta}$ and the regressive adjustment to the tax schedule.

In sum, under these conditions, the presence of a third sector makes the optimal income tax schedule more regressive compared to a standard two-sector economy.

5.2 Increasing or Decreasing Returns to Scale

Taking $N = 2$, now consider any homothetic production function $Y(E) = h(\tilde{Y}(E))$, where $h(\tilde{Y})$ is some increasing function with elasticity $\varepsilon_h(E) \equiv h'(\tilde{Y}(E))\tilde{Y}(E)/Y(E)$, and $\tilde{Y}(E)$ has CRS as in the preceding subsection, with substitution elasticity $\sigma(\rho)$. Suppose the total output Y is divided across sectors according to the \tilde{Y} -income shares, i.e. $r_1(E)E_1 \equiv Y_1(E) = \alpha(\rho)Y(E)$ and $r_2(E)E_2 \equiv Y_2(E) = (1 - \alpha(\rho))Y(E)$, where $\rho = E_1/E_2$ and $\alpha(\rho) =$

$\tilde{Y}_1(\rho)E_1 / (\tilde{Y}_1(\rho)E_1 + \tilde{Y}_2(\rho)E_2)$. Using the $\bar{\xi}$ defined in (42), Lemma 8 yields the following:

Proposition 5. *Suppose $N = 2$, $Y(E) = h(\tilde{Y}(E))$ with $\tilde{Y}(E)$ linear homogeneous, and with returns $r_i(E) = \tilde{Y}_i(E_1, E_2)Y(E) / \tilde{Y}(E)$. Then the optimal correction factor in (12) is*

$$1 - \sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i(E)} = 1 + \left(\frac{f_E^1(w)}{f_E(w)} - \alpha(\rho) \right) \bar{\xi} - (1 - \varepsilon_h(E)). \quad (44)$$

The optimal adjustment in (44) can be transparently decomposed into two terms: a local correction exactly as in Corollary 1, and a global correction $1 - \varepsilon_h(E)$ which uniformly scales up or down marginal keep shares $1 - T'(y)$. In particular, if $\varepsilon_h(E) < (>)1$, we have decreasing (increasing) returns to scale and marginal tax rates are scaled up (down) relative to an economy with CRS.

Note that unless $\varepsilon_h = 1$, the direction of no relative return effects in E -space, $(\rho, 1)$, and the direction of zero externalities, $(-1/r_1, 1/r_2)$, are always distinct. By Proposition 2, β has full rank and both consistency constraints are needed in this example.

5.3 A Pure Resource Transfer Activity

The preceding subsection allowed for aggregate externalities but fixed the sectoral composition of incomes at the aggregate level via the CRS income shares $\alpha(\rho)$ and $1 - \alpha(\rho)$. Keeping $N = 2$, we now consider the opposite case where $Y(E)$ exhibits CRS but the activity 1 income share $a(E)$ is not equal to $\alpha(\rho)$, so that one activity is underpaid—and the other overpaid—relative to its marginal product.

Consider in particular the extreme example where $Y(E) = E_2$ but $a(E) = a(E_1)$ is a positive and increasing function. Then activity 1 is pure “stealing” of (or getting credit for) output, which is produced exclusively in activity 2, as discussed in the introduction.²² Because activity 1 is purely extractive, the Pigouvian tax is $\tau_p^1 = 1$, whereas activity 2 generates positive externalities by increasing the returns $r_1 = a(E_1)E_2/E_1$ to activity 1 and therefore commands a Pigouvian subsidy $\tau_p^2 = -a/(1 - a)$. In contrast to the preceding section, externalities are purely distributional; consequently, $a\tau_p^1 + (1 - a)\tau_p^2 = 0$, i.e., there is no Pigouvian correction in aggregate.

Defining the elasticity of a as $\varepsilon_1(E_1) \equiv (\partial a(E_1) / \partial E_1)E_1 / a(E_1)$ and using the definitions of C and S from Corollary 1 yields the following result:

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

²²In Appendix D, we provide general formulas for less extreme cases, with qualitatively similar results.

$$\sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\bar{\zeta}_i}{r_i} = \left(\tau_p^1 - (1 - a - \varepsilon_1) \bar{\zeta} \right) \frac{f_E^1(w)}{f_E(w)} + \left(\tau_p^2 + a \bar{\zeta} \right) \frac{f_E^2(w)}{f_E(w)} \quad (45)$$

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

The numerator of $\bar{\zeta}$ is positive if the rent-seeking activity 1 is also the high income activity—since then $I_1 > 0$ if incentive constraints are down-binding, and $R_1 \geq 0$ if Pareto weights are (weakly) higher among same-wage earners with higher income shares in the productive activity 2.²³

The terms in parentheses in (45) are sums of the Pigouvian tax rates and relative return effect adjustments. The latter are intuitive. For example, a subsidy on activity 2 raises E_2 and thereby increases the relative returns to activity 1, which leads to a wasteful effort shift towards activity 1. The optimum therefore involves an *undercorrection* relative to the Pigouvian subsidy, as reflected by $a\bar{\zeta}$ in the second term of (45).

The relative return adjustment in activity 1 is ambiguous: it depends on $\varepsilon_1 \gtrless 1 - a$. This is because an increase in E_1 has two offsetting effects: first, by increasing $a(E_1)$, it increases the relative returns to activity 1. Second, it causes *crowding* in activity 1: the earnings $a(E_1)E_2$ are spread over a larger effort E_1 . If $\varepsilon_1 < 1 - a$, then the latter effect dominates, and taxes on activity 1 cause perverse flows of effort towards activity 1. The optimal correction on activity 1-intensive parts of the income distribution is therefore below the Pigouvian one when $\varepsilon_1 < 1 - a$ (and above it when $\varepsilon_1 > 1 - a$).

As in the preceding example, this problem requires both consistency constraints, since the no-externality direction $(a/r_1(E), (1 - a)/r_2(E))$ does not coincide with the direction $(a/r_1(E), (1 - a - \varepsilon_1)/r_2(E))$ of zero relative return effects.

5.4 Externalities from One Activity

Next, suppose that returns depend only on aggregate effort in one activity, i.e., $r_i(E) = r_i(E_1)$ for all $i = 1, \dots, N$. Rothschild and Scheuer (2014) is a special case, with $N = 2$, in which a rent-seeking activity 1 imposes negative externalities on both activities, so $\beta_1^j < 0$, while activity two imposes no externalities, so $\beta_2^j = 0$. We treat here the case with general N and general externalities generated by sector 1.

In particular, since $\beta_i^j = 0$ for all $i = 2, \dots, N$ and all j , the matrix β has rank one and, per Proposition 3, the problem can be written with a single consistency constraint—namely the constraint for E_1 . This leads to the following result:

²³Lemma 9 in Appendix C shows that the denominator—and hence $\bar{\zeta}$ here—is positive when a natural stability condition is met.

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

$$\bar{\xi} = \frac{\tau_p^1 + \sum_{j=1}^{N-1} \Delta\beta_1^j (I_j + R_j) / r_1}{1 - \sum_{j=1}^{N-1} \Delta\beta_1^j (C_{j1} + S_{j1})}. \quad (46)$$

The corrective factor $\bar{\xi}$ is weighted by the local income share of activity 1 and deviates from the Pigouvian correction τ_p^1 only if there are relative return effects—i.e., if $\Delta\beta_1^j \neq 0$ for some j . These effects enter in an intuitive way. For instance, suppose activity 1 generates negative externalities, so $\tau_p^1 > 0$. Then the denominator in (46) increases $\bar{\xi}$ relative to τ_p^1 if an increase in E_1 on average raises the relative returns to activities j with $C_{j1}, S_{j1} > 0$, and vice versa. In this case, the increase in E_1 indirectly causes a reinforcing flow of effort into activity 1. Conversely, a tax on sector 1 income directly and beneficially reduces E_1 and indirectly leads to effort flows that further reduce E_1 . Such a tax is therefore even more desirable than based on the purely Pigouvian motives.²⁴

The second term in the numerator of (46) 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 (i.e., if $\Delta\beta_1^j$ is positively correlated with I_j, R_j).²⁵ Then an increase in the marginal income tax at wage levels where activity 1 is prevalent lowers E_1 and indirectly redistributes by raising returns to lower-wage, high redistributive preference activities. Of course, analogous results can be obtained from (46) when the tax leads to the opposite sectoral shifts or when activity 1 imposes positive or mixed externalities.

For $N = 2$ we obtain the special case in Rothschild and Scheuer (2014), with

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

and C and S given by (37) and (38). If $l(w)$ is increasing, so that $C > 0$, and if we also have $I_1, R_1 > 0$ (because the externality-causing activity 1 is also a high wage and low redistributive preference activity), then an over- (under-)correction with $\bar{\xi} > (<) \tau_p^1$ is optimal if and only if $\Delta\beta_1^1 > (<) 0$. Our general framework shows that these results hold for arbitrary forms of externalities generated by activity 1 (rather than just negative ones).

²⁴Again, the denominator of (46) is positive if the optimum is stable per Lemma 9 in Appendix C.

²⁵Since $\sum_j I_j = \sum_j R_j = 0$, $\sum_j \Delta\beta_1^j (I_j + R_j)$ is N times the across- j covariance $\text{Cov}(\Delta\beta_1^j, I_j + R_j)$.

5.5 Externalities Targeted at One Activity

Finally, consider the case where $r_1(E)$ is general but $r_i(E) = r_i$ are constants for all $i = 2, \dots, N$ —so that only the first activity *bears* any externalities. A simple example with $N = 2$ is another specification of a pure resource transfer activity, with $Y(E) = Y(E_1)$ and $Y_1(E) = Y(E_1) - E_2$ and $Y_2(E_2) = E_2$. Here, all output is produced through activity 1, and activity 2 takes away some of this output one-for-one, as discussed in Section 2.2.

Generally, $\beta_i^j = 0$ for all $j \neq 1$ and β again has rank one in this case, this time with all *columns* being zero except for the first, which has elements β_i^1 (and $\Delta\beta = \beta$). Intuitively, any movement in E -space that changes $r_1(E)$ generates both an externality and a relative return change. Conversely, since $t_p^i = -\beta_i^1 Y_1$ in this example, all the $N - 1$ dimensions of \mathbb{R}^N orthogonal to the vector $(\beta_1^1, \beta_2^1, \dots, \beta_N^1)$ are directions of both no externalities and no relative return effects because changes of E in these directions leave $r_1(E)$ unchanged. Per Proposition 3, we need only one consistency constraint in the outer problem—a β_i^1 -weighted sum of the original N constraints (8).

By (34), $\xi_i / \beta_i^1 = \xi_1 / \beta_1^1$ for all i , which yields the following result:

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

$$\sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = \sum_{i=1}^N \frac{f_E^i(w)}{f_E(w)} \frac{\beta_i^1}{r_i} \bar{\xi} \quad \text{with} \quad \bar{\xi} = \frac{-Y_1 + I_1 + R_1}{1 - \sum_{i=1}^N \beta_i^1 (C_{1i} + S_{1i})}. \quad (47)$$

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., when $\sum_i (f_E^i(w) / f_E(w)) (\beta_i^1 / r_i) = 0$. Intuitively, this is a wage level at which a variation in the marginal income tax rate leads to changes in E that leaves $r_1(E)$ unaffected, so the optimal marginal tax rate is “as if” all returns were fixed locally.

Otherwise, the Y_1 term in the numerator of $\bar{\xi}$ in (47) captures the Pigouvian taxes on all activities that affect r_1 (since $\tau_p^i = \beta_i^1 Y_1 / r_i$). 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 these taxes; the intuition for these terms is the same as in the earlier examples.

6 Conclusion

We have developed a general framework for tax policy that applies to a wide range of imperfect labor markets with rich heterogeneity and general equilibrium effects—effects

which we believe are ubiquitous. As complementary work has shown, our tools can be productively operationalized for quantitative optimal tax analyses in important special cases. Most recently, Ales, Murnaz and Sleet (2014) have performed income tax simulations based on a N -sector model similar to ours but with more restrictive assumptions on individual heterogeneity and the economy-wide production function that allow them to quantify the impact of technical progress between the 1970s and the 2000s on optimal tax progressivity. In Rothschild and Scheuer (2013, 2014), we have computed optimal policies for the special case with $N = 2$ and where no or only one activity generates externalities, and shown how to empirically identify the underlying two-dimensional skill-distribution under some assumptions. These studies demonstrate that general equilibrium effects can be of considerable quantitative importance. We expect the flexible toolkit developed here to be useful for further quantitative studies under less restrictive assumptions.

Our analysis also provides general insights into the type of empirical evidence needed for calculating optimal taxes in the presence of externalities. For instance, in the pure resource transfer example discussed in Section 5.3, the Pigouvian component of the correction would be entirely pinned down by the aggregate income share accruing to the transfer activity. The optimal correction depends additionally on the elasticity of this income share with respect to effort in the transfer activity: high elasticities imply an optimal correction strictly greater than the Pigouvian component; low elasticities imply a sub-Pigouvian correction. Thus, information on these income shares and elasticities would be of direct use for optimal policy design. More generally, a high-level policy lesson is that evidence on the source and magnitude of externalities are insufficient for designing tax policy; information on where the externalities are borne is also critical.

Finally, we expect our methods to be useful for policy design in other important contexts, including general equilibrium effects from consumption externalities, status concerns (through “keeping up with the Joneses” preferences or conspicuous consumption), moral hazard in insurance, hidden savings and side trades, household labor supply, and multitask problems in team production. We leave these for future research.

References

- ALES, L., M. KURNAZ, AND C. SLEET (2014): “Tasks, Talents, and Taxes,” *Mimeo, Tepper School of Business, Carnegie Mellon University*.
- ATKINSON, A., T. PIKETTY, AND E. SAEZ (2011): “Top Incomes in the Long-Run of History,” *Journal of Economic Literature*, 49, 3–71.

- BESLEY, T., AND M. GHATAK (2013): "Bailouts and the Optimal Taxation of Bonus Pay," *American Economic Review, Papers and Proceedings*, 103, 163–167.
- BIAIS, B., T. FOUCAULT, AND S. MOINAS (2011): "Equilibrium High Frequency Trading," *Mimeo, TSE*.
- BOOTHBY, W. (1986): *An Introduction to Differentiable Manifolds and Riemannian Geometry*. Academic Press, Orlando, Florida, 2 edn.
- CHONÉ, P., AND G. LAROQUE (2010): "Negative Marginal Tax Rates and Heterogeneity," *American Economic Review*, 100, 2532–2547.
- CLARKE, F. (1976): "A New Approach to Lagrange Multipliers," *Mathematics of Operations Research*, 1, 165–174.
- CREMER, H., F. GAHVARI, AND N. LADOUX (1998): "Externalities and Optimal Income Taxation," *Journal of Public Economics*, 70, 343–364.
- DIAMOND, P. (1973): "Consumption Externalities and Imperfect Corrective Pricing," *Bell Journal of Economics and Management Science*, 4, 526–538.
- DIAMOND, P. (1998): "Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Tax Rates," *American Economic Review*, 88, 83–95.
- DIXIT, A. (1985): "Tax Policy in Open Economies," in *Handbook of Public Economics*, ed. by A. Auerbach, and M. Feldstein, pp. 2475–2490. North Holland, Amsterdam.
- FUDENBERG, D., AND J. TIROLE (1991): *Game Theory*. MIT Press, Cambridge.
- GLODE, V., AND R. LOWERY (2012): "Compensating Financial Experts," *Mimeo, University of Pennsylvania*.
- GOLOSOV, M., A. TSYVINSKI, AND N. WERQUIN (2014): "A Variational Approach to the Analysis of Tax Systems," *Mimeo, Princeton University and Yale University*.
- GOMES, R., J.-M. LOZACHMEUR, AND A. PAVAN (2014): "Differential Taxation and Occupational Choice," *Mimeo, Toulouse School of Economics and Northwestern University*.
- GUESNERIE, R. (1981): "On taxation and incentives: Further reflections on the limits to redistribution," *Discussion paper 89, Inst. für Ges.- und Wirtschaftswiss., Wirtschaftstheoretische Abt., Univ. Bonn*.
- HAMMOND, P. J. (1979): "Straightforward individual incentive compatibility in large economies," *The Review of Economic Studies*, pp. 263–282.
- HELLWIG, M. (2007): "The Undesirability of Randomized Income Taxation under Decreasing Risk Aversion," *Journal of Public Economics*, 91, 791–816.
- HENDREN, N. (2014): "The Inequality Deflator: Interpersonal Comparisons without a Social Welfare Function," *NBER Working Paper No. 20351*.

- JACQUET, L., AND E. LEHMANN (2014): "Optimal Nonlinear Income Taxation with Multidimensional Types: The Case with Heterogeneous Behavioral Responses," *Thema Working Paper No. 2014-01*.
- JACQUET, L., B. VAN DER LINDEN, AND E. LEHMANN (2013): "The Optimal Marginal Tax Rate with both Extensive and Intensive Responses," *Journal of Economic Theory*, 148, 1770–1805.
- KLEVEN, H. J., C. T. KREINER, AND E. SAEZ (2009): "The Optimal Income Taxation of Couples," *Econometrica*, 77, 537–560.
- KOPCZUK, W. (2003): "A Note on Optimal Taxation in the Presence of Externalities," *Economics Letters*, 80, 81–86.
- LOCKWOOD, B., C. NATHANSON, AND G. WEYL (2014): "Taxation and the Allocation of Talent," *Mimeo, Harvard University and University of Chicago*.
- LOCKWOOD, B., AND M. WEINZIERL (2014): "De Gustibus Non Est Taxandum: Theory and Evidence on Preference Heterogeneity and Redistribution," *Mimeo, Harvard University*.
- LUENBERGER, D. G. (1969): *Optimization by vector space methods*. John Wiley & Sons.
- MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.
- MYERSON, R. B. (1979): "Incentive compatibility and the bargaining problem," *Econometrica*, pp. 61–73.
- NAITO, H. (1999): "Re-Examination of Uniform Commodity Taxes Under A Non-Linear Income Tax System And Its Implication for Production Efficiency," *Journal of Public Economics*, 71, 165–188.
- PIKETTY, T., E. SAEZ, AND S. STANTCHEVA (2014): "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 6, 230–271.
- ROTHSCHILD, C., AND F. SCHEUER (2013): "Redistributive Taxation in the Roy Model," *Quarterly Journal of Economics*, 128, 623–668.
- (2014): "Optimal Taxation with Rent-Seeking," *Mimeo, Wellesley College and Stanford University*.
- ROY, A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.
- SADKA, E. (1978): "On the Optimal Taxation of Consumption Externalities," *Quarterly Journal of Economics*, 92.
- SAEZ, E. (2001): "Using Elasticities to Derive Optimal Tax Rates," *Review of Economic Studies*, 68, 205–229.

SANDMO, A. (1975): “Optimal taxation in the presence of externalities,” *The Swedish Journal of Economics*, 77, 86–98.

SCHEUER, F. (2014): “Entrepreneurial Taxation with Endogenous Entry,” *American Economic Journal: Economic Policy*, 6, 126–163.

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

WERNING, I. (2007): “Pareto-Efficient Income Taxation,” *Mimeo, MIT*.

A Appendix: Proofs

A.1 Proof of Lemma 1

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

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

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

A.2 Proof of Lemma 3

It suffices to show that $\int_w h(w) dF_{E^n}^i(w)$ converges to $\int h(w) dF_E^i(w)$ for any continuous function h such that $\sup_w |h(w)| \leq \bar{h}$ for some \bar{h} . Using the definition of F_E^i and changing variables yields

$$\begin{aligned} \int_w h(w) dF_{E^n}^i(w) &= \int_w \int_{\Phi} h(w) q_{E^n}^i(\varphi) dG_{E^n}(w, \varphi) \\ &= \int_{\Theta} h(w_{E^n}(\theta)) q_{E^n}^i(\theta) dF(\theta). \end{aligned} \quad (48)$$

The sequence $h(w_{E^n}(\theta)) q_{E^n}^i(\theta)$ is bounded by \bar{h} and converges pointwise to $h(w_E(\theta)) q_E^i(\theta)$ by Lemma 2. The dominated convergence theorem implies $\int_w h(w) dF_{E^n}^i(w) \rightarrow \int_{\Theta} h(w_E(\theta)) q_E^i(\theta) dF(\theta) = \int_w h(w) dF_E^i(w)$.

A.3 Proof of Lemma 4

The proof closely follows Appendix B in Hellwig (2007), which considers a problem with changing, weakly convergent wage distributions. This lemma considers a fixed E , and therefore a fixed type distributions, so much of his apparatus is unnecessary here. We employ this apparatus because the same arguments will apply to our proof of Lemma 5. To keep the discussion parallel to Hellwig’s proofs, we work with allocations described by $c(w)$ and $y(w)$ instead of $V(w)$ and $l(w)$. The idea is to construct a feasible allocation $(\hat{c}(w), \hat{y}(w))$ which achieves the supremum in (6).

Start with a sequence $(c^k(w), y^k(w))$ of feasible allocations for which $\int u(c^k(w), y^k(w)/w) d\Psi_E(w) \geq W(E) - 2^{-k}$. By a diagonalization argument (viz Hellwig’s Lemma B.3), and because c and y are, by assumption, uniformly bounded, we can find a subsequence, indexed by k_n , of allocations which converges

pointwise to some $(\bar{c}(w), \bar{y}(w))$ for each rational w . By (7), each of the allocations $(c^k(w), y^k(w))$, and therefore the (partial) allocation $(\bar{c}(w), \bar{y}(w))$ are monotonic in w . The allocation $(\hat{c}(w), \hat{y}(w))$:

$$(\hat{c}(w), \hat{y}(w)) \equiv \lim_{w' \downarrow w, w \in \mathbb{Q}} (\bar{c}(w'), \bar{y}(w')) \quad (49)$$

is therefore well defined. Per the arguments in Hellwig's Lemma B.4, $(\hat{c}(w), \hat{y}(w))$ satisfies (7). Per the arguments in Hellwig's Lemma B.5, and the fact that the type distributions F_E^i and Ψ_E are independent of k , $(\hat{c}(w), \hat{y}(w))$ satisfies (8) and (9), and has

$$\int u(\hat{c}(w), \hat{y}(w)/w) d\Psi_E(w) = \lim_{n \rightarrow \infty} \int u(c^{k_n}(w), y^{k_n}(w)/w) d\Psi_E(w) \geq \lim_{n \rightarrow \infty} W(E) - 2^{-k_n} = W(E). \quad (50)$$

A.4 Proof of Lemma 5

Take any sequence $E^k \rightarrow E$ with a convergent limit $\bar{W} = \lim_{k \rightarrow \infty} W(E^k)$. For each k , let $\{c^k(w), y^k(w)\}$ be a feasible and optimal allocation, which exists by Lemma 4. As in the proof of Lemma 4 above, we can find a subsequence, indexed by k_n , which converges pointwise to some $(\bar{c}(w), \bar{y}(w))$ for each rational w , and then define

$$(\hat{c}(w), \hat{y}(w)) \equiv \lim_{w' \downarrow w, w \in \mathbb{Q}} (\bar{c}(w'), \bar{y}(w')). \quad (51)$$

The allocation $(\hat{c}(w), \hat{y}(w))$ so defined satisfies (7). Per Lemma 3, F_E^i , F_E and Ψ_E are weakly continuous in E . Lemma B.1 in Hellwig (2007) therefore implies that $(\hat{c}(w), \hat{y}(w))$ also satisfies (8) and (9) (with the wage distributions at E). It also implies $\int u(\hat{c}(w), \hat{y}(w)/w) d\Psi_E(w) = \lim_{n \rightarrow \infty} W(E^{k_n}) = \bar{W}$. Hence, $W(E) \geq \bar{W}$, completing the proof.

A.5 Proof of Proposition 1

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

$$\begin{aligned} \mathcal{L} &= \int_{\underline{w}_E}^{\bar{w}_E} V(w) \psi_E(w) dw - \int_{\underline{w}_E}^{\bar{w}_E} V(w) \hat{\eta}'(w) \lambda dw + \int_{\underline{w}_E}^{\bar{w}_E} u_l(c(V(w), l(w)), l(w)) \frac{l(w)}{w} \hat{\eta}(w) \lambda dw \\ &+ \sum_{i=1}^N \xi_i \lambda \left[E_i - \frac{1}{r_i(E)} \int_{\underline{w}_E}^{\bar{w}_E} w l(w) f_E^i(w) dw \right] + \lambda \int_{\underline{w}_E}^{\bar{w}_E} (w l(w) - c(V(w), l(w))) f_E(w) dw. \end{aligned} \quad (52)$$

This Lagrangian approach is valid (i.e. constraint qualification holds) generically by Theorem 3 in Clarke (1976). Using $\partial c / \partial V = 1/u_c$ and compressing notation, the first order condition for $V(w)$ is

$$\hat{\eta}'(w) \lambda = \psi_E(w) - \lambda f_E(w) \frac{1}{u_c(w)} + \hat{\eta}(w) \lambda \frac{u_{cl}(w) l'(w)}{u_c(w) w}. \quad (53)$$

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

$$\eta'(w) = \psi_E(w) \frac{u_c(w)}{\lambda} - f_E(w) + \eta(w) \frac{u_{cc}(w) c'(w) + u_{cl}(w) l'(w) + u_{cl}(w) l(w)/w}{u_c(w)}. \quad (54)$$

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

$$u_c(w)c'(w) + u_l(w)l'(w) + u_l(w)\frac{l(w)}{w} = 0, \quad (55)$$

the fraction in (54) 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 (54) and rearranging yields

$$-\frac{\partial MRS(w)}{\partial c}l(w)\frac{y'(w)}{y(w)}\eta(w) = f_E(w) - \psi_E(w)\frac{u_c(w)}{\lambda} + \eta'(w). \quad (56)$$

Integrating this ODE gives

$$\begin{aligned} \eta(w) &= \int_w^{\bar{w}_E} \left(f_E(s) - \psi_E(s)\frac{u_c(s)}{\lambda} \right) \exp \left(\int_w^s \frac{\partial MRS(t)}{\partial c} l(t) \frac{y'(t)}{y(t)} dt \right) ds \\ &= \int_w^{\bar{w}_E} \left(1 - \frac{\psi_E(s)}{f_E(s)} \frac{u_c(s)}{\lambda} \right) \exp \left(\int_w^s \left(1 - \frac{\varepsilon^u(t)}{\varepsilon^c(t)} \right) \frac{dy(t)}{y(t)} \right) f_E(s) ds, \end{aligned} \quad (57)$$

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

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

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

which after some algebra can be rewritten as

$$w f_E(w) \left(1 - \frac{MRS(w)}{w} \right) - w \sum_{i=1}^N \frac{\xi_i}{r_i(E)} f_E^i(w) = \eta(w) \left(\frac{\partial MRS(w)}{\partial l} \frac{l}{w} + \frac{MRS(w)}{w} \right), \quad (58)$$

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

$$1 - \sum_{i=1}^N \frac{\xi_i}{r_i(E)} \frac{f_E^i(w)}{f_E(w)} = (1 - T'(y(w))) \left[1 + \frac{\eta(w)}{w f_E(w)} \left(1 + \frac{\partial MRS(w)}{\partial l} \frac{l}{MRS(w)} \right) \right]. \quad (59)$$

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 (57) and (59).

A.6 Proof of Lemma 6

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

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

because $\sum_{j=1}^N \text{Cov}(q_E^j, q_E^k | w) = \text{Cov}(\sum_{j=1}^N q_E^j, q_E^k | w) = \text{Cov}(1, q_E^k | w) = 0$ for all w . For S_{kj} , we prove the result by showing that $\sum_{j=1}^N Q_k^j(x_E(\phi)) = 0$ for all $\phi \in \Phi$. For this, use (32) and (33) to write

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

Hence, showing that $\sum_j \Omega_j \partial Z_j / \partial(\theta_k r_k(E)) = 0$ will complete the proof. Using (32), we have

$$Z_j(x_E(\phi)) \equiv \frac{r_j(E)\theta_j}{w} = r_j(E)\theta_j \min_{p \in \Delta^{N-1}} m\left(\frac{p_1}{r_1(E)\theta_1}, \dots, \frac{p_N}{r_N(E)\theta_N}\right), \quad (60)$$

so

$$\frac{\partial Z_j}{\partial(r_k \theta_k)} = \frac{\delta_{kj}}{w} - r_j \theta_j m_k(e/y) \frac{q^k}{(r_k \theta_k)^2}, \quad (61)$$

where m_k denotes the (homogeneous of degree zero) partial derivative of m w.r.t. its k -th argument. Note that the first order conditions for the minimization in (60) are $m_k(e/y)/(r_k \theta_k) = m_N(e/y)/(r_N \theta_N)$ for all $k = 1, \dots, N$, which implies

$$\frac{1}{w} = \frac{m(e)}{y} = m(e/y) = \sum_{k=1}^N m_k(e/y) \frac{q^k}{r_k \theta_k} = \frac{m_N(e/y)}{r_N \theta_N} \sum_{k=1}^N q^k = \frac{m_N(e/y)}{r_N \theta_N} = \frac{m_k(e/y)}{r_k \theta_k} \quad \forall k, \quad (62)$$

where the second equality uses linear homogeneity of m , the third uses Euler's theorem, and the fourth and sixth the first order conditions. Substituting this in (61) and using $\Omega_j = e_j/m(e) = e_j w/y$ yields

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

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

(ii) $\sum_{k=1}^N S_{kj}(E) = 0$ follows from (27) and $\sum_{k=1}^N Q_k^j(x_E(\phi)) x_E^k(\phi) = 0$ for all ϕ . To see the latter, use (33) and

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

by the zero-homogeneity of ζ and Euler's theorem. $\sum_{k=1}^N C_{kj}(E) = 0$ follows from (25) and an analogous argument to part (i).

A.7 Proof of Proposition 2

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

A.8 Proof of Proposition 3

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

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

A.9 Proof of Lemma 8

Dropping the arguments E , the optimality conditions (34) can be written for $N = 2$ as

$$\mathbf{A}\vec{\xi} = \vec{t}_p + \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} (I_1 + R_1). \quad (63)$$

Since $(\Delta\beta_1^1, \Delta\beta_2^1)'$ is an eigenvector of \mathbf{A} , it is also an eigenvector of \mathbf{A}^{-1} (with associated eigenvalue $1/\gamma_2$), and we can write (63) as

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

Moreover, defining the eigenbasis

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

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

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

Hence,

$$\vec{\xi} = \vec{t}_p + \begin{pmatrix} \Delta\beta_1^1 \\ \Delta\beta_2^1 \end{pmatrix} \frac{I_1 + R_1 - b (\Delta\beta_2^1/r_2 - \Delta\beta_1^1/r_1) (C + S)}{\gamma_2}. \quad (64)$$

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

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

Substituting in (64) yields the result.

A.10 Proof of Proposition 4

Observe first that $\beta_i^3 = 0$ for all i , so the third row of $\Delta\beta$ is zero. Using this together with $\vec{t}_p = 0$ in (34) immediately implies $\xi_3 = 0$. Straightforward calculations yield $\beta_1^1(E) = -\beta_2^1(E)/\rho = Y_1'(\rho)/(E_2Y_1(\rho))$, $\beta_1^2(E) = -\beta_2^2(E)/\rho = Y_2'(\rho)/(E_2Y_2(\rho))$, $\Delta\beta_1^1(E) = -1/(E_1\hat{\sigma}(\rho))$, and $\Delta\beta_2^1(E) = 1/(E_2\hat{\sigma}(\rho))$, where $-1/\hat{\sigma}(\rho) = \rho(Y_1'(\rho)/Y_1(\rho) - Y_2'(\rho)/Y_2(\rho))$ is the substitution elasticity of \hat{Y} . Hence, the second row of $\Delta\beta$ is $-\rho$ times the first row, which implies $\xi_2 = -\rho\xi_1$. We can therefore use the first row of the system (34) to solve for ξ_1 , which (using $\beta_1^3 = 0$) yields $\xi_1/r_1 = -(1 - \hat{\alpha})\hat{\xi}$ where $\hat{\xi}$ is given in (41), and $\xi_2/r_2 = -\rho\xi_1/r_1 = \hat{\alpha}\hat{\xi}$. Finally, substituting these two and $\xi_3 = 0$ in the adjustment term delivers (40).

A.11 Proof of Proposition 5

Tedious algebra yields

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

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

so $\Delta\beta_1^1(\rho)E_1 = -\Delta\beta_2^2(\rho)E_2 = -1/\sigma(\rho)$. Moreover, $\tau_p^1(E) = \tau_p^2(E) = 1 - \varepsilon_h(E)$. Substituting in (36) yields $\xi_1/r_1 = 1 - \varepsilon_h(E) - (1 - \alpha(\rho))\bar{\xi}$ and $\xi_2/r_2 = 1 - \varepsilon_h(E) + \alpha(\rho)\bar{\xi}$, and using this in (12) yields (44).

A.12 Proof of Proposition 6

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

A.13 Proof of Proposition 7

We can use the first row of (34) to explicitly solve for ξ_1 , using $\xi_i = 0$ for all $i \neq 1$:

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

where $t_p^1 = -\sum_{j=1}^N \beta_1^j Y^j$ and C_{j1} and S_{j1} are given in (25) and (27). This immediately yields the result.

A.14 Proof of Proposition 8

Using $\xi_i = \xi_1 \beta_i^1 / \beta_1^1$, we can use the first equation in the system (34) to solve for ξ_1 :

$$\xi_1 = \left(t_p^1 + \beta_1^1(I_1 + R_1) \right) / \left(1 - \sum_{i=1}^N \beta_i^1(C_{i1} + S_{i1}) \right)$$

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

B Differentiability of $W(E)$

In this appendix, we provide a simple but standard example where differentiability of $W(E)$ can be directly established. Consider quasilinear and iso-elastic preferences $u(c, l) = c - l^{1+1/\varepsilon} / (1 + 1/\varepsilon)$. Then $\eta(w) = \Psi_E(w) - F_E(w)$ and $\lambda = 1$. Combining the wage w -worker's first order condition with (12) yields

$$l(w)^{1/\varepsilon} = w(1 - T'(y(w))) = w \left(1 - \sum_{j=1}^N \frac{\zeta_j}{r_j(E)} \frac{f_E^j(w)}{f_E(w)} \right) \left[1 + \frac{\Psi_E(w) - F_E(w)}{w f_E(w)} \left(1 + \frac{1}{\varepsilon} \right) \right]^{-1}. \quad (65)$$

Substituting (65) into the consistency constraints (8), denoting by $X_E(w)$ the term in square brackets above and considering the case with $\varepsilon = 1$ yields

$$r_i(E)E_i = \int_{\underline{w}_E}^{\bar{w}_E} w^2 \left(f_E^i(w) - \sum_{j=1}^N \frac{\zeta_j}{r_j(E)} \frac{f_E^i(w) f_E^j(w)}{f_E(w)} \right) X_E(w)^{-1} dw. \quad (66)$$

Equivalently, in matrix notation, $A(E)\hat{\xi} = b(E)$, where

$$A_{ij}(E) = \int_{\underline{w}_E}^{\bar{w}_E} w^2 \frac{f_E^i(w) f_E^j(w)}{f_E(w)} X_E(w)^{-1} dw, \quad b_i(E) = \int_{\underline{w}_E}^{\bar{w}_E} w^2 f_E^i(w) X_E(w)^{-1} dw - r_i(E)E_i, \quad \text{and } \hat{\xi}_j = \zeta_j / r_j(E).$$

If $A(E)$ is invertible at E , this system has a unique solution $\hat{\xi}^*$, which is continuous in E . The optimal allocation $(y(w), c(w))$ is therefore unique and continuous in E . The envelope theorem then implies that the partial derivatives of $W(E)$ exist and coincide with the partial derivatives of the Lagrangian, as assumed in Section 3.4. Since the space of singular (square) matrices is a lower-dimensional sub-manifold of the space of square matrices, the approach used in Section 3.4 is valid for almost all E in almost all problems.²⁶

C Eigenvalues and Stability

Holding the schedule $l(w)$ fixed, the right-hand sides of the system of consistency constraints (8) defines a mapping $E \rightarrow \mathcal{E}(E)$. An optimal E is obviously a fixed point of this mapping, and it is reasonable to assume it is a stable fixed point, 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 3.4, (34) involves varying the schedule $l(w)$ to $\tilde{l}_E(w)$, with corresponding mapping $E \rightarrow \hat{\mathcal{E}}(E)$, where $\hat{\mathcal{E}}(E)$ is the vector with elements $\hat{\mathcal{E}}_i(E) = \int_{\underline{w}_E}^{\bar{w}_E} w \tilde{l}_E(w) f_E^i(w) dw / r_i(E)$. Stability of a fixed point of this mapping requires that all eigenvalues of the Jacobian of the dynamic system $\dot{E}_i = \hat{\mathcal{E}}_i(E) - E_i$, $i = 1, \dots, N$, have negative real parts. By our derivation of the consistency constraint effects in section 3.4, this Jacobian is equal to $-A$. We therefore have:

Lemma 9. *A fixed point E of the mapping $E \rightarrow \hat{\mathcal{E}}(E)$ defined above is stable if and only if all eigenvalues of the matrix $A = I_N - \Delta\beta(C + S)$ in (34) have positive real parts.*

²⁶Moreover, if $A(E)$ is singular, almost all $b(E)$ -vectors will be such that there is no $\hat{\xi}$ that solves $A(E)\hat{\xi} = b(E)$ —i.e., there will be no optimal allocation at such E s.

D General Sectoral Income Shares

Let $Y(E)$ have constant returns to scale and $Y_1(E) = a(E)Y(E)$ and $Y_2(E) = (1 - a(E))Y(E)$, so that

$$r_1(E) = a(E)Y(E)/E_1 \quad \text{and} \quad r_2(E) = (1 - a(E))Y(E)/E_2. \quad (67)$$

Defining $\varepsilon_1(E) \equiv \frac{\partial a(E)}{\partial E_1} \frac{E_1}{a(E)}$ and $\varepsilon_2(E) \equiv \frac{\partial(1-a(E))}{\partial E_2} \frac{E_2}{1-a(E)}$ yields:

Proposition 9. *If $N = 2$, $Y(E)$ has constant returns to scale and private returns are given by (67), then the adjustment to the marginal tax rate formula in (12) is*

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

$$\bar{\xi} = \frac{I_1 + R_1 + \frac{a-\alpha}{a(1-a)}(C+S)}{a(1-a)Y + (1-\varepsilon_1-\varepsilon_2)(C+S)}. \quad (68)$$

The proof of Proposition (9) involves straightforward algebraic computations. The first terms in the two brackets are simply the Pigouvian corrections for the two activities, since $\tau_p^1 = (a - \alpha)/a$ and $\tau_p^2 = (\alpha - a)/(1 - a)$, weighted by 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 $\tau_p^2 < 0$ (and as in Section 5.3, $a\tau_p^1 + (1 - a)\tau_p^2 = 0$). The terms multiplied by $\bar{\xi}$ in the brackets capture the deviations from the Pigouvian corrections due to the relative return effects of a variation in the marginal tax rate. The intuition, which relies on the crowding effects captured by the elasticities ε_1 and ε_2 , is identical to Section 5.3.

The formula is particularly transparent in the case where a is homogenous of degree zero, so $a(\rho)$ with $\rho = E_1/E_2$.

Corollary 2. *If $a(E)$ is homogeneous of degree zero (and $\bar{\xi}$ is given in (68)), then*

$$\sum_{i=1}^2 \frac{f_E^i(w)}{f_E(w)} \frac{\xi_i}{r_i} = \frac{1}{1 - a} \left(\frac{f_E^1(w)}{f_E(w)} - a \right) \left(\frac{a - \alpha}{a} - (1 - a - \varepsilon_1)\bar{\xi} \right). \quad (69)$$

The first bracketed term, which parallels the corresponding terms in (42) and (44), 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.