# Optimal Taxation with Rent-Seeking*

**Casey Rothschild**
Wellesley College

**Florian Scheuer**
Stanford University and NBER

Revised Version
April 2012

**Abstract**

We develop a framework for optimal income taxation when agents can choose between working in a traditional sector, where private and social products coincide, and a rent-seeking sector, where the private returns to effort exceed the social returns. This could result from the fact that income in this sector reflects the capture of preexisting rents, or because rent-seekers reduce the returns to traditional work (within-versus cross-sector externalities). We characterize Pareto optimal non-linear income taxes when the government does not observe whether an individual is a traditional worker or a rent-seeker. We identify a sectoral shift effect as a key determinant for optimal tax policy. If the within-sector externality dominates, it can blunt the incentive to tax the highest wage earners at very high rates, even if they are socially unproductive rent-seekers and the government has an intrinsic desire for progressive redistribution. Intuitively, taxing their effort at a lower rate stimulates their rent-seeking efforts, thereby congesting the rent-seeking sector and discouraging further entry.

*Email addresses: crothsch@wellesley.edu, scheuer@stanford.edu. We are grateful to the editor, Philippe Jehiel, three anonymous referees, and Daron Acemoglu, Alan Auerbach, Marco Bassetto, Michael Boskin, Paco Buera, V. V. Chari, Peter Diamond, Emmanuel Farhi, Caroline Hoxby, Patrick Kehoe, Chris Phelan, Thomas Piketty, James Poterba, Emmanuel Saez, Matthew Weinzierl, Iván Werning, David Wildasin and seminar participants at Alberta, Berkeley, Cornell, Guelph, the Federal Reserve Bank of Minneapolis, Konstanz, Munich, MIT, Stanford, UMass-Amherst, Uppsala, Wellesley, Wharton, ETH Zurich, SED Ghent, the NBER Summer Institute and Public Economics Program Meeting, the Minnesota Workshop in Macroeconomic Theory, the NTA Annual Conference on Taxation 2011 and the ASSA Meetings 2012 for numerous valuable suggestions. All errors are our own.

# 1 Introduction

The unwinding of the financial crisis over the past several years has exposed numerous examples of highly compensated individuals whose apparent contributions to social output proved illusory. Events like the recent housing bubble provide fertile ground for rent-seeking: pursuing personal enrichment by extracting a slice of the existing economic pie rather than by increasing the size of that pie. These highly salient examples of rent-seeking activities have inspired calls for a more steeply progressive tax code. For instance, Paul Krugman argued for higher taxes on "supersized incomes" in the context of discussing the profits from high speed trading, on the grounds that "it is hard to see how traders who place their orders one-thirtieth of a second faster than anyone else do anything to improve that social function."[1] Moreover, in various countries, the introduction of very high taxes (up to 90%) on bonus payments in the financial sector has been discussed on the grounds of similar rent-seeking arguments.

The argument behind such proposals is intuitively appealing. If much of the economic activity at high incomes is primarily socially unproductive rent-seeking or "skimming," then it would seem natural for a well designed income tax code to impose high marginal rates at high income levels.[2] This would discourage such behavior while simultaneously raising revenues which could be used, e.g., to lower taxes and encourage more productive effort at lower income levels.

This paper adapts the Mirrlees (1971) framework and provides a formal foundation for studying the implications of such rent-seeking activities for optimal income taxation. We use this framework to characterize optimal income taxes in the presence of a broad class of rent-seeking externalities and to provide formal conditions under which this simple intuition is or is not valid.

By way of illustration, consider a simple model with two sectors.[3] The productive sector has output proportional to (skill-weighted) aggregate effort. In the other sector, workers compete for a fixed rent $\bar{\mu} > 1$. There is an equal measure of two types of workers, with preferences $u(c,e) = c - e^{\gamma}/\gamma$ over consumption and effort. Type 1 workers produce one unit of output per unit of effort in the productive sector. Type 2 workers have zero productivity in that sector. Wages in the rent-seeking sector are equal to $\bar{\mu}/E$ and $\varphi_R\bar{\mu}/E$ for type 1 and type 2 workers, respectively, where $E = \lambda e_1 + \varphi_R e_2$ is the (skill weighted) aggregate rent-seeking effort and $\lambda$ is the fraction of type 1 workers who work in the rent-seeking sector. These rent-seeking wages correspond to a situation in which

[1] New York Times Editorial, August 3, 2009.

[2] See Bertrand and Mullainathan (2001) for evidence of such rents.

[3] We thank one of the referees for suggesting this example.

each unit of equivalent effort claims an equal share of the total rent $\bar{\mu}$, and one unit of type 2's effort is equivalent to $\varphi_R > 1$ units of type 1's effort.

It is straightforward to show that utilitarian social welfare is maximized with *zero* taxes whenever $\overline{\mu} - 1 < \varphi_R^{\gamma/(\gamma-1)} < \overline{\mu}$ and $\gamma$ is sufficiently large. To wit, there are three possibilities: Either $E < \bar{\mu}$, and all type 1 workers are rent-seekers, $E > \bar{\mu}$, and they all work in the traditional sector, or $E = \bar{\mu}$, and some fraction $\lambda \in [0,1]$ works in rent-seeking. In this third case, the wages of the two types are $w_1 = 1$ and $w_2 = \varphi_R$, total income is $e_1 + \varphi_R e_2$, and utilitarian social welfare is $W = e_1 + \varphi_R e_2 - \left(e_1^\gamma + e_2^\gamma\right)/\gamma$. The zero-tax efforts $e_i^{\gamma-1} = w_i$, $i = 1, 2$, clearly lead to maximal social welfare, $W^*$, among these $E = \bar{\mu}$ allocations, and they are consistent with $\lambda \in [0,1]$ whenever $\varphi_R^{\gamma/(\gamma-1)} \in [\overline{\mu} - 1, \overline{\mu}]$. For large enough $\gamma$, $W^* = (1 + \varphi_R^{\gamma/(\gamma-1)})(\gamma - 1)/\gamma$ exceeds the social welfare of $\bar{\mu}$ obtained in the allocation with 100 percent taxation and $E = 0$—which is the only other candidate for an optimum, since if $\bar{\mu} \neq E > 0$, decreasing the effort $e_i$ of any type who works in the rent-seeking sector has no effect on output and is thus welfare enhancing.

Rent-seekers in this example are clearly identifiable and their efforts produce no output, so that the Pigouvian tax on their effort would be 100%. Yet zero taxes are optimal. To see why, consider imposing a small tax on them, reducing their effort by $\delta$. This decrease in total rent-seeking effort $E$ raises the returns to rent-seeking $\overline{\mu}/E$. Productive workers then shift into rent-seeking until $E = \overline{\mu}$ is restored. The net effect is an income reduction of exactly $w_2\delta$. Although rent-seekers are not directly productive, their *indirect productivity* is therefore exactly equal to their wage: by congesting the rent-seeking sector, they help to maintain the productivity of the type 1 workers in the productive sector.

We call this indirect productivity the *sectoral shift effect*. The optimality of *zero*-taxes is a knife-edge feature of this stylized example. It provides a stark illustration of the importance of general equilibrium effects from occupational choice for optimal taxation in the presence rent-seeking. We show that the importance of these effects is robust by developing a framework with a continuum of types and an arbitrary two-dimensional distribution of skills for traditional and rent-seeking activities. In this framework, rent-seeking effort may impose externalities both on other rent-seekers—as in the preceding example—and/or across-sector externalities on productive workers.

We then characterize the set of income taxes that maximize some weighted average of the individuals' utilities in the economy under the assumption that taxes cannot or do not condition on sector of employment. They turn out to be characterized by a multiplicative correction to standard optimal tax formulas for economies without rent-seeking. Whenever the within-sector externalities are larger than the across sector externalities, this correction is strictly below the Pigouvian correction, which would align private and social marginal products of effort. This is because of a generalized sectoral shift effect:[4] raising marginal taxes in portions of the income distribution with rent-seekers discourages the externality causing activity there. But in so doing, it raises the relative returns to rent-seeking activity. This causes a sectoral shift of workers into rent-seeking, reducing the net benefits of the direct externality reduction.

When the across-sector rent-seeking externalities dominate the within-sector externalities, the sectoral shift effect reverses sign, and the optimal correction *exceeds* the Pigouvian correction. Only in the knife-edge case where the within and across-sector externalities exactly balance—or when all workers are rent-seekers—does the naive Pigouvian correction apply. This provides a transparent qualification for whether rent-seeking leads to higher or more progressive taxes: does an additional unit of rent-seeking effort, keeping everyone else's behavior fixed, increase or decrease the relative returns to rent-seeking?

Solving for the optimal income taxes in our framework is a more challenging problem than a standard Mirrlees (1971) optimal tax problem. The first reason is a fixed point problem, which results from the endogeneity of the wage distribution. Fixing any given tax code, the decision of an individual about which sector to work in depends on the wages she can earn in the rent-seeking and traditional sectors. The former depends on how much effort other individuals are exerting in the rent-seeking sector. Solving for the outcomes induced by that tax code thus involves finding a level of aggregate rent-seeking effort $E$ such that the wages induced by $E$ lead to sectoral choices and effort such that aggregate rent-seeking effort is indeed $E$.

The second challenge is that we employ a two-dimensional distribution of skill-types, for which standard techniques typically do not apply (see Rochet and Choné, 1998). We address it by observing that the realized wage distribution is well defined conditional on any given aggregate rent-seeking effort $E$. Since taxes depend only on income, a standard single-crossing property holds. This allows us to treat the problem as a fixed point problem for $E$ nested within a Mirrleesian optimal income tax problem.

Proposition 1 provides a partial characterization of the solution to this problem. It reveals that the optimal marginal keep shares (one minus the marginal tax rate) can, as alluded to above, be expressed as a standard optimal tax formula multiplied by an extra correction term. This structure is consistent with the "principle of targeting" (Dixit, 1985) and, more specifically, the "additivity principle" discussed in the expansive literature on corrective taxation in the presence of atmospheric externalities, according to which taxes can be expressed as a sum of the optimal Pigouvian taxes and the optimal taxes from a

[4] As we show in the following, the sectoral shift effect is reinforced by several other effects which arise with overlapping and continuous sectoral wage distributions.

related problem without externalities.[5] In fact, when there is *no* traditional sector and all workers are rent-seekers, our problem can be fit into Kopczuk's (2003) framework and the optimal correction is precisely the Pigouvian tax that aligns the private and social marginal products of effort (see Proposition 2). Since the correction is uniform, there is no change in the optimal progressivity of the tax schedule due to rent-seeking.

In the general two-sector case, the correction term in the optimal tax formula diverges from this Pigouvian tax. Proposition 3 shows that this divergence depends in a transparent way on the direction of the sectoral shift effect. This result is most closely related to Diamond (1973), although our motivation, framework, and corrective instruments are quite distinct (see the discussion in section 3.6). He shows that the optimal linear tax on an externality producing consumption good can be expressed as a Pigouvian correction that captures the *direct* effect of the tax on the demand for the good, and an adjustment term that reflects the indirect, general equilibrium effect of the changes in consumption of the good induced by the direct effect. We have a distinct motivation for such general equilibrium effects, namely in the form of occupational choice, which has not received attention is this literature so far. This allows us to provide more specific insights into how the optimal correction should deviate from the Pigouvian tax rate. Moreover, with rent-seeking, the externality arises from income, which is qualitatively different from the consumption externalities that this literature has focused on.[6]

While rent-seeking is a conceptually important element of our model, our methods more closely track the optimal income taxation literature, notably Mirrlees (1971), Diamond (1998), Saez (2001) and Werning (2007). Our paper also contributes to recent efforts to study optimal taxation under multidimensional private heterogeneity. In a recent study of the optimal income taxation of couples, Kleven, Kreiner and Saez (2009) have made progress along these lines, as have Choné and Laroque (2010). Both papers have significantly different information structures than ours, however. The second dimension of heterogeneity enters preferences additively in the former, and in the latter it is a taste for labor rather than a second standard skill type as we employ here.

We build on work in the rent-seeking literature, starting from Tullock (1967). Our model of rent-seeking is broad enough to include a wide range of activities, such as the patent races discussed in Dixit (1987), socially useless but privately profitable financial

[5] See Sandmo (1975), Sadka (1978), Bovenberg and van der Ploeg (1994), Pirttilä and Tuomala (1997), Cremer, Gahvari, and Ladoux (1998), Kopczuk (2003), and the discussion in section 3.5.

[6] Bovenberg and Goulder (1996) characterize optimal corrective taxation with a production externality and consider general equilibrium effects from the fact that a pollution tax affects the real wage and therefore labor supply and tax revenue. However, this effect results from the restriction to a linear labor tax and would disappear in their model without heterogeneity with a nonlinear income tax that we consider here.

speculation discussed by Arrow (1973), and a broad class of externality causing activities.

Finally, our paper relates to the literature studying the equilibrium allocation of talent across different sectors when there are rents to be captured in some sectors (see e.g. Baumol, 1990, Acemoglu and Verdier, 1998 and Cahuc and Challe, 2009). None of these papers considers optimal tax policy to correct these equilibrium outcomes, which is the focus of our contribution. An important exception is the recent work by Philippon (2010), who considers an endogenous growth model with financiers, workers and entrepreneurs. He analyzes the effect of linear, sector specific taxes on growth. In contrast, we do not consider growth effects of taxation nor sector-specific tax instruments.

Our paper proceeds as follows. Section 2 describes our modeling framework. Section 3 studies optimal non-linear taxation. Proposition 1 provides a partial characterization of the optimal marginal taxes. It shows that the marginal keep shares can be written as the product of a standard optimal tax formula and a correction factor that arises from the rent-seeking externality. Proposition 2 shows that this correction factor is equal to $1 - t_{Pigou}$ in a one-sector economy in which all workers are rent-seekers, where $t_{Pigou}$ is the wedge between the private and social marginal returns to rent-seeking effort. Proposition 3 shows that the correction factor diverges from $t_{Pigou}$ in a generic two-sector model because of the sectoral shift and sympathetic effects. Section 4 concludes. All proofs appear in a technical appendix.

## 2 The Model

We consider an economy with two sectors: A traditional sector, where private and social marginal products coincide, and a rent-seeking sector, where the private marginal product exceeds the social marginal product. There is a unit-measure continuum of individuals who can choose to work in either sector. Each individual is endowed with a two-dimensional skill vector $(\theta, \varphi) \in \Theta \times \Phi$, $\Theta = [\underline{\theta}, \overline{\theta}]$, $\Phi = [\underline{\varphi}, \overline{\varphi}]$, where $\theta$ captures an individual's skill in the traditional sector (the "$\Theta$-sector"), and $\varphi$ captures her skill in the rent-seeking sector (the "$\Phi$-sector"). Skills are distributed with cdf $F : \Theta \times \Phi \rightarrow [0, 1]$ and associated pdf $f(\theta, \varphi)$. Preferences are characterized by a continuously differentiable and concave utility function $u(c, e)$ defined over consumption $c$ and effort $e$ with $u_c > 0$, $u_e < 0$ and $u_{cc}, u_{ee} \leq 0$.

Since the income tax is not sector specific, each individual chooses the sector she works in so as to maximize her wage, which depends on both her skill and the wage per unit of effective effort in that sector.[7] To describe the latter, let aggregate output be given by

$$Y(E_\theta, E_\varphi) = Y_\theta + Y_\varphi = E_\theta \Gamma(E_\varphi) + \mu(E_\varphi), \quad \text{where}$$

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

are the aggregate effective effort in the traditional and rent-seeking sector, respectively, $e(\theta,\varphi)$ is the effort of type $(\theta,\varphi)$, and $P(E_\varphi)$, defined below, is the set of types working in the rent-seeking sector. The wage of a $\Theta$-sector worker of skill type $(\theta,\varphi)$ is given by $w_\theta = \theta\Gamma(E_\varphi)$. We assume $\Gamma'(E_\varphi) \leq 0$, so that rent-seeking can have a negative effect on output and wages in the traditional sector. The same type's rent-seeking wage is given by $w_\varphi = \varphi\mu(E_\varphi)/E_\varphi$, where $\mu(E_\varphi)$ is total output in the rent-seeking sector with $\mu(0) = 0$, $\mu'(E_\varphi) \geq 0$, $\mu''(E_\varphi) < 0$. For any $E_\varphi$, wage maximization implies

$$w = \max\left\{\theta\Gamma(E_\varphi), \varphi\frac{\mu(E_\varphi)}{E_\varphi}\right\} \quad \text{and} \quad P(E_\varphi) \equiv \left\{(\theta,\varphi) \in \Theta\times\Phi \,\middle|\, \theta\Gamma(E_\varphi) < \varphi\frac{\mu(E_\varphi)}{E_\varphi}\right\}.$$

Since both wages and occupational choice depend only on aggregate rent-seeking effort, we simplify notation by dropping the subscript and letting $E = E_\varphi$ henceforth.

Observe that wages in the traditional sector reflect the social marginal product of effort in that sector. In contrast, wages diverge from the social marginal product of effort in the rent-seeking sector for two reasons. First, rent-seeking may impose a negative *cross-sectoral* externality $\Gamma'(E) \leq 0$ on traditional workers. Second, each rent-seeker's share $e \times \varphi\mu(E)/E$ of the total income $\mu(E)$ earned in the rent-seeking sector reflects their share of the aggregate rent-seeking effort $\varphi e/E$—that is, rent-seeking wages reflect the average within-sector product of labor. Since $\mu(E)$ displays decreasing returns, rent-seeking effort imposes a negative *within-sector* externality $\mu'(E) - \mu(E)/E < 0$. In other words, the average product $\mu(E)/E$ that individuals face as their wage exceeds the social marginal product of effort $\mu'(E)$. One extreme case would arise if $\Gamma(E) \equiv 1$ and $\mu(E) \equiv E$, so that the rent-seeking problem disappears. On the other hand, "pure" rent-seeking occurs when $\mu(E) = \overline{\mu}$ so that there is a fixed rent to be captured in the rent-seeking sector and any effort there is in fact completely unproductive since $\mu'(E) = 0$, as in the example in the introduction.

Our setup captures a fully general two-sector model of negative externalities in which (i) only one of the two sectors imposes externalities and (ii) each increment of effort in a

[7] Note that our results trivially extend to the case where individuals can work in both sectors, but preferences are given by $u(c, e_\theta + e_\varphi)$, where $e_\theta$ and $e_\varphi$ denote effort exterted in the $\Theta$- and $\Phi$-sector, respectively.

given sector is equally remunerated.[8] Its close connection with standard models of rent-seeking is illustrated by the following two examples:

**Example 1** (Contests). *Consider $N$ rent-seekers competing for a rent of value $\bar{\mu}$. As in Tullock (1980), the probability $p_i$ that player $i \in \{1, ..., N\}$ wins the contest is*

$$p_i(\varphi_i e_i, \varphi_{-i} e_{-i}) = \frac{\varphi_i e_i}{\sum_{j=1}^{N} \varphi_j e_j}.$$

*Player i's expected payoff is therefore $\varphi_i e_i \overline{\mu}/E$ with $E \equiv \sum_{j=1}^{N} \varphi_j e_j$. Whenever $\varphi_i e_i/E$ is small, the private marginal returns to individual rent-seeking effort are given by $\varphi_i \overline{\mu}/E$, as in our general model, and exceed the zero social marginal returns to effort.*

**Example 2** (Races). *Suppose individuals race to discover a rent with value $M(t)$ at time $t$. The first individual to discover it captures the entire benefit. Let the hazard rate for individual $i$ to find a so far undiscovered rent be given by $\lambda \varphi_i e_i$, $\lambda > 0$. The probability that* some *rent-seeker discovers it in a time interval $dt$ is then simply $\lambda E dt$, where, again $E = \sum_i \varphi_i e_i$. The time to discovery therefore follows an exponential distribution with $p(t|E) = \lambda E \exp(-\lambda E t)$, and the expected payoff to an individual rent-seeker $i$ is*

$$\frac{\varphi_i e_i}{E} \int_0^\infty M(t) p(t|E) dt \equiv \varphi_i e_i \frac{\mu(E)}{E} \quad \textit{with} \quad \mu(E) \equiv \int_0^\infty M(t) p(t|E) dt.$$

As demonstrated by these examples, our framework is flexible enough to address the examples of rent-seeking discussed in the introduction, such as high-speed trading or arbitrage seeking, where a winner-takes-all compensation can induce a divergence between privately faced returns and social marginal product.

# 3 Optimal Non-linear Income Taxation

## 3.1 Definitions and Preliminaries

We use cumulative Pareto weights $\Psi(\theta, \varphi)$ with the corresponding density $\psi(\theta, \varphi)$ to obtain Pareto efficient allocations. The social planner maximizes $\int_{\Theta \times \Phi} V(\theta, \varphi) d\Psi(\theta, \varphi)$ subject to resource and self-selection constraints, where $V(\theta, \varphi)$ is the utility of agents of type $(\theta, \varphi)$. The observation that makes the social planner's problem tractable is that fixing $E$

[8] To see this, note that $Y_\theta = \Gamma(E) E_\theta$, since the income earned in the sector without externalities must be proportional to $E_\theta$ given (ii). Then define $Y_\varphi = Y(E_\theta, E) - \Gamma(E) E_\theta \equiv \mu(E)$. Since wages are proportional to $\mu(E)/E$ in the externality-causing sector, negative within-sector externalities (i.e., the only externalities when $E_\theta = 0$) require $\mu(E)/E > \mu'(E)$ for all $E$, which implies $\mu''(E) < 0$.

determines the wage of each worker. For any given $E$, we can therefore define the cumulative wage distributions $F_E(w) \equiv F(w, wE/\mu(E))$ and $\Psi_E(w) \equiv \Psi(w, wE/\mu(E))$ with associated densities $f_E(w)$ and $\psi_E(w)$.

We are particularly interested in studying the "regular" case in which the social planner assigns greater welfare weight to low-wage individuals, as in the following definition:

**Definition 1.** *Welfare weights are* regular *if $\psi_E(w)/f_E(w)$ is non-increasing in $w$ for any $E$.*

We also refer to Pareto optima that correspond to such welfare weights as regular allocations. For some of our analysis, we will focus on a particularly compelling sub-set of the Pareto optimal allocations which result from redistributive motives across different wage-earners, but not across individuals with the same wage yet in different sectors. Such allocations can be obtained using *relative Pareto weights*, characterized by a non-decreasing function $\tilde{\Psi} : [0,1] \to [0,1]$. This function is used to define cumulative Pareto weights via $\Psi(\theta, \varphi) = \tilde{\Psi}(F(\theta, \varphi))$. The associated density over wages for given $E$ is $\tilde{\psi}_E(w) = \tilde{\Psi}'(F_E(w)) f_E(w) \geq 0$, so relative welfare weights are regular whenever $\tilde{\Psi}$ is weakly concave. In particular, this implies that $\tilde{\Psi}(F_E(w)) \geq F_E(w)$ for all $w$, so that the social planner puts higher weight on lower quantiles of the wage distribution than the population shares, resulting in a desire to (weakly) redistribute from higher to lower wage individuals.[9]

Since fixing $E$ fixes a wage distribution, it reduces the optimal income tax problem to a one-dimensional screening problem despite the underlying two-dimensional heterogeneity in the population. In particular, allocations can only condition on an individual's wage $w$, so that we can write $c(w)$, $e(w)$, $y(w) \equiv we(w)$ and $V(w) \equiv u(c(w), e(w))$ for the consumption, effort, income and utility of an individual with wage $w$. As is standard, we assume that preferences satisfy single-crossing, i.e. $-u_e(c, y/w)/(u_c(c, y/w)w)$ is decreasing in $w$, so that an individual's marginal rate of substitution between income and consumption is decreasing in the $(y, c)$-space. We also denote the uncompensated and compensated wage elasticities of effort as a function of the wage by $\varepsilon^u(w)$ and $\varepsilon^c(w)$, respectively. Note that $\varepsilon^c(w) \geq 0$ and single-crossing implies that $\varepsilon^u(w) > -1$. For some of our examples, we will consider the specific form of quasilinear and isoelastic preferences, so that $u(c, e) = c - e^{1+1/\varepsilon}/(1 + 1/\varepsilon)$, where both the uncompensated and compensated wage elasticity of effort are constant and equal to $\varepsilon > 0$.

[9] We show in Lemma 3 in Section 3.4 that relative welfare weights imply that, for any $E$, the average welfare weights of individuals who earn the same wage but in different sectors are the same, i.e. $\psi^\theta_E(w)/f^\theta_E(w) = \psi^\varphi_E(w)/f^\varphi_E(w)$ for all $w$. Hence, in this case the planner does not care about redistributing across individuals in different sectors per se, but only across different wage earners.

### 3.2 A Decomposition and Pareto Optimality

It is useful to decompose the problem of finding Pareto optimal allocations into two steps. The first (referred to as "inner problem") involves finding the optimal resource-feasible and incentive-compatible allocation for a fixed level of rent-seeking effort $E$ and thus a fixed wage distribution with $F_E(w) = F\left(w/\Gamma(E), wE/\mu(E)\right)$,

$$f_E^\theta(w) = \frac{1}{\Gamma(E)} \int_{\underline{\varphi}}^{w\frac{E}{\mu(E)}} f\left(\frac{w}{\Gamma(E)}, \varphi\right) d\varphi, \quad f_E^\varphi(w) = \frac{E}{\mu(E)} \int_{\underline{\theta}}^{w} f\left(\theta, w\frac{E}{\mu(E)}\right) d\theta$$

and $f_E(w) = f_E^\theta(w) + f_E^\varphi(w)$, and analogous Pareto weights $\Psi_E(w)$ and corresponding densities. Since $E$ also determines the sectoral choice of all individuals, we call $f_E^\theta(w)$ and $f_E^\varphi(w)$ the densities of wages in the traditional and rent-seeking sectors, respectively. We denote the support of the wage distribution for given $E$ by $[\underline{w}_E, \overline{w}_E]$ with the lowest and highest wages $\underline{w}_E = \max\left\{\underline{\theta}\Gamma(E), \underline{\varphi}\mu(E)/E\right\}$ and $\overline{w}_E = \max\left\{\overline{\theta}\Gamma(E), \overline{\varphi}\mu(E)/E\right\}$. The second step then involves finding the optimal level of $E$. We call this the "outer" problem.

Since the wage distribution is fixed for given $E$, the inner problem for the Pareto optimum is an almost standard Mirrlees problem with the only complication that we have to take into account the sectoral composition of the economy. More precisely, the induced level of equivalent effort in the rent-seeking sector has to be consistent with the level of $E$ that we started from. For some given Pareto weights $\Psi(\theta, \varphi)$ and hence induced Pareto weights $\Psi_E(w)$, we therefore define the inner problem as follows (where $c(V, e)$ is the inverse function of $u(c, e)$ w.r.t. the first argument):

$$W(E) \equiv \max_{V(w), e(w)} \int_{\underline{w}_E}^{\overline{w}_E} V(w) d\Psi_E(w) \tag{1}$$

s.t.

$$V'(w) + u_e(c(V(w), e(w)), e(w))\frac{e(w)}{w} = 0 \;\; \forall w \in [\underline{w}_E, \overline{w}_E] \tag{2}$$

$$\mu(E) - \int_{\underline{w}_E}^{\overline{w}_E} we(w) f_E^\varphi(w) dw = 0 \tag{3}$$

$$\int_{\underline{w}_E}^{\overline{w}_E} we(w) f_E(w) dw - \int_{\underline{w}_E}^{\overline{w}_E} c(V(w), e(w)) f_E(w) dw \geq 0. \tag{4}$$

We employ the standard Mirrleesian approach of optimizing directly over allocations, i.e. over effort $e(w)$ and consumption or, equivalently, utility $V(w)$ profiles. The social planner then maximizes a weighted average of the individuals' utilities $V(w)$ subject to a set of constraints. (4) is a standard resource constraint and constraint (3) guarantees that total effort in the rent-seeking sector indeed sums up to $E$ (or, equivalently, the sum of all incomes in the rent-seeking sector equals $\mu(E)$). Finally, the allocation $V(w), e(w)$ needs to be incentive compatible, i.e.

$$V(w) \equiv u(c(w), e(w)) = \max_{w'} u\left(c(w'), \frac{e(w')w'}{w}\right). \tag{5}$$

It is a well-known result that under single-crossing, the global incentive constraints (5) are equivalent to the local incentive constraints (2) and the monotonicity constraint that income $y(w)$ must be non-decreasing in $w$.[10] We follow the standard approach of dropping the monotonicity constraint and checking ex-post that it is satisfied. If the solution to problem (1) to (4) does not satisfy it, optimal bunching would need to be considered.

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

## 3.3 Marginal Tax Rate Formulas from the Inner Problem

Based on solving the inner problem (1) to (4) for given $E$, we obtain the following formula for optimal marginal tax rates in any Pareto optimum:

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

$$1 - T'(y(w)) = \left(1 - \xi \frac{f_E^{\varphi}(w)}{f_E(w)}\right)\left(1 + \frac{\eta(w)}{w f_E(w)} \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)}\right)^{-1} \tag{6}$$

*with*

$$\eta(w) = \int_w^{\overline{w}_E} \left(1 - \frac{\psi_E(x)}{f_E(x)} \frac{u_c(x)}{\lambda}\right) \exp\left(\int_w^x \left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)}\right) \frac{dy(s)}{y(s)}\right) f_E(x) dx \tag{7}$$

*for all $w \in [\underline{w}_E, \overline{w}_E]$, where $\lambda$ is the multiplier on the resource constraint (4) and $\lambda\xi$ the multiplier on the consistency constraint (3).*

These formulas are the same as those for a standard Mirrlees model (see e.g. equations (15) to (17) in Saez, 2001), with the only difference that, at each wage, marginal keep shares $1 - T'(y(w))$ are scaled down by the correction factor $1 - \xi f_E^{\varphi}(w)/f_E(w)$, where $\xi$ is the Lagrangian on constraint (3) and $f_E^{\varphi}(w)/f_E(w)$ is the share of rent-seekers at wage level $w$. This is intuitive as it is saying that the optimal correction, which makes agents internalize the rent-seeking externality, is proportional to the fraction of rent-seekers at $w$ and the shadow cost of the consistency constraint (3). As usual, the term $\eta(w)$ captures the

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

redistributive motives of the government and income effects from the terms in the exponential function. A particularly transparent formula can be obtained under the assumption of quasilinear preferences, so that income effects disappear. Then $u_c(w) = \lambda = 1$ and $\varepsilon^u(w) = \varepsilon^c(w)\ \forall w$, so that the marginal tax rate simplifies as follows:

**Corollary 1.** *With quasilinear preferences, the marginal tax rate in any Pareto optimum without bunching is such that*

$$1 - T'(y(w)) = \left(1 - \xi \frac{f_E^{\varphi}(w)}{f_E(w)}\right)\left(1 + \frac{\Psi_E(w) - F_E(w)}{w f_E(w)}\left(1 + \frac{1}{\varepsilon(w)}\right)\right)^{-1}. \tag{8}$$

In this case, $T'(y(w)) \geq 0$ at all income levels if and only if $\Psi_E(w) \geq F_E(w)$ (which is implied if Pareto weights are regular), and it is increasing in $\Psi_E(w) - F_E(w)$, i.e. in the degree to which $\Psi_E(w)$ shifts weight to lower wage individuals compared to $F_E(w)$. This captures the redistributive effect of an increase in the marginal tax rate at $w$. Moreover, $T'(y(w))$ is decreasing in the wage elasticity of effort $\varepsilon(w)$ and the wage density $f_E(w)$, which are both related to the distortionary effects at $w$ (see also Diamond, 1998). Again, the corrective motives for income taxation are captured by the term $1 - \xi f_E^{\varphi}(w)/f_E(w)$, which scales up marginal tax rates in proportion to the share of rent-seekers.[11]

Under any preference assumptions, the top marginal tax rate is given by $T'(y(\overline{w}_E)) = \xi f_E^{\varphi}(\overline{w}_E)/f_E(\overline{w}_E)$, and by $\xi$ if the share of rent-seekers at the top is one. We next consider the outer problem in more detail in order to explore the determination of $E$ and $\xi$.

## 3.4 Optimal Size of the Rent-Seeking Sector from the Outer Problem

We start with the following useful result about $\xi$ from the outer problem:

**Lemma 1.** $\xi > 0$ *in any Pareto optimum with regular Pareto weights.*

Hence, the correction factor in the marginal keep share formula (6) is always smaller than one under regular redistributive motives, so that the optimal income tax schedule

[11] In Rothschild and Scheuer (2011), we make explicit this comparison between a Pareto optimal tax schedule as characterized by (8) and the tax schedule that would be optimal in a standard Mirrlees problem, using the notion of a self-confirming policy equilibrium. I.e., rather than comparing optimal tax formulas for different economies, we ask which tax schedule would be set in the present economy by a government that ignores the endogeneity of wages due to rent-seeking. Such a government would set marginal tax rates according to the same formula as in a standard Mirrlees model (i.e. as in (8) but without the correction factor). However, for the equilibrium to be self-confirming, i.e. consistent with constraint (3), $E$ would need to solve the fixed point problem $E = \tilde{E}(E)$, where $\tilde{E}(E)$ is the level of aggregate rent-seeking effort induced by the solution to a standard Mirrlees problem with wage distribution $F_E(w)$.

indeed scales up marginal tax rates whenever there are rent-seekers. In particular, this implies that the top marginal tax rate is positive in any regular Pareto optimum.

Our main goal is to compare $\xi$ to the "naive" Pigouvian tax, which is defined by

$$\left(1 - t_{Pigou}\right) \frac{\mu(E)}{E} \equiv \mu'(E) + \Gamma'(E) E_\theta,$$

i.e. as the tax that aligns the private and social returns to rent-seeking effort. This tax would make agents internalize the damage from an additional unit of effective rent-seeking effort, holding everyone else's behavior fixed. The key question in the following will be how the correct externality correction $\xi$, taking all the general equilibrium effects (in particular from occupational choice) into account, differs from the intuitive, partial equilibrium one, as captured by $t_{Pigou}$.[12] For this purpose, it is useful to define the elasticity of wages with respect to $E$ in each sector as

$$\beta^\theta(E) \equiv -\Gamma'(E)\frac{E}{\Gamma(E)} > 0 \text{ and } \beta^\varphi(E) \equiv -\frac{d}{dE}\left(\frac{\mu(E)}{E}\right)\frac{E}{\mu(E)/E} = 1 - \frac{\mu'(E)E}{\mu(E)} > 0.$$

Using these definitions allows us to express the Pigouvian tax rate as

$$t_{Pigou} = \beta^\varphi(E) + \frac{Y_\theta}{Y_\varphi}\beta^\theta(E) > 0,$$

which accounts for the effect of $E$ on wages in the rent-seeking sector (where $\beta_\varphi(E)$ captures the within-sectoral externality) and the traditional sector (captured by the output-weighted cross-sectoral externality $\beta_\theta(E)$). Let $\Delta\beta(E) \equiv \beta^\varphi(E) - \beta^\theta(E)$ indicate the relative importance of the within- versus across-sectoral externality. Using this, Lemma 2 provides a decomposition of the welfare effect of marginal changes in $E$:

**Lemma 2.** *At any Pareto optimum without bunching, the welfare effect of a marginal change in aggregate rent-seeking effort $E$ is*

$$W'(E) = -\lambda\frac{\mu(E)}{E}t_{Pigou} + \frac{\Delta\beta(E)}{E}(I+R) + \xi\lambda\left(\frac{\mu(E)}{E} + \frac{\Delta\beta(E)}{E}(C+S)\right), \tag{9}$$

*where*

$$I \equiv \lambda\int_{\underline{w}_E}^{\overline{w}_E} \eta(w) w \frac{V'(w)}{u_c(w)}\frac{d}{dw}\left(\frac{f_E^\varphi(w)}{f_E(w)}\right)dw, \tag{10}$$

[12] This definition of the Pigouvian tax is in line with Diamond (1973) and Cremer, Gahvari, and Ladoux (1998). Kopczuk (2003) defines the Pigouvian tax directly as implicitly taking all general equilibrium effects into account, corresponding to our $\xi$, but that definition would prevent us from getting at the question we are after.

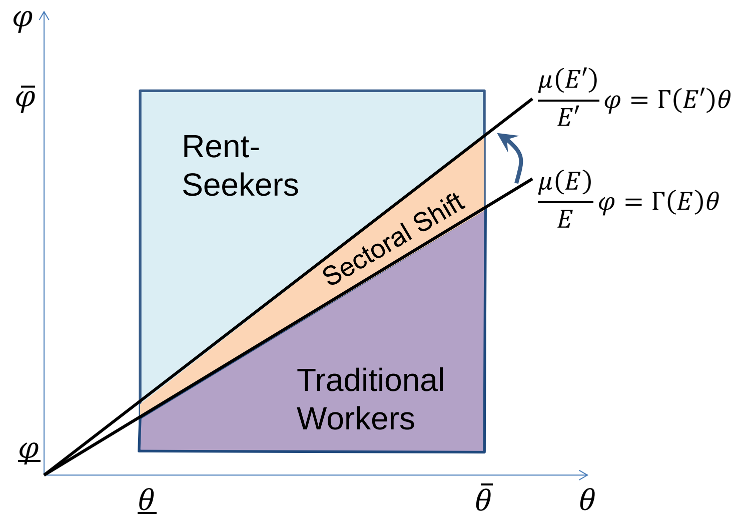


Figure 1: The sectoral shift effect if $\Delta\beta(E) > 0$, $E' > E$.

$$R \equiv \int_{\underline{w}_E}^{\overline{w}_E} V'(w) w \frac{f_E^\theta(w) f_E^\varphi(w)}{f_E(w)} \left( \frac{\psi_E^\theta(w)}{f_E^\theta(w)} - \frac{\psi_E^\varphi(w)}{f_E^\varphi(w)} \right) dw, \tag{11}$$

$$C \equiv \int_{\underline{w}_E}^{\overline{w}_E} w^2 e'(w) \frac{f_E^\theta(w) f_E^\varphi(w)}{f_E(w)} dw \tag{12}$$

*and*

$$S \equiv \int_{\underline{w}_E}^{\overline{w}_E} w^2 e(w) f\left( \frac{w}{\Gamma(E)}, \frac{wE}{\mu(E)} \right) dw. \tag{13}$$

The terms with $-\lambda(\mu(E)/E)t_{Pigou}$ and $\xi\lambda\mu(E)/E$ in (9) come from the direct effect of changes in $E$ on (i) the consistency condition (3) and (ii) on wages $w$ in (3) and (4). If the terms described in (10) to (13) terms were zero, then (14) would imply $\xi = t_{Pigou}$, as in the one-sector example we consider below. In the general case, these terms capture the *indirect* effect of $E$ through (i) changes in sectoral choice (term $S$) and (ii) the effect of changing wages on $V(w)$ and $e(w)$ (terms $I$, $R$, and $C$).

Figure 1 illustrates the sectoral shift effect $S$. When $E$ rises, wages in both sectors fall. When $\Delta\beta(E) > 0$ and the within sector externality dominates the cross-sectoral one, *relative* wages rise in the traditional sector, inducing some individuals to shift out of rent-seeking.

The terms $C$ and $R$ arise from the fact that an increase in $E$ has different effects on individuals who originally had the same wage but worked in different sectors. For example, when $\Delta\beta(E) > 0$, individuals in the traditional sector shift *up* along the $V$ and $e$ schedules relative to rent-seekers, effectively re-allocating $V$ and $e$ across individuals who were at the same original wage. When $e'(w) > 0$, this causes an additional net shift of effort out of rent-seeking, and $C$ (which captures this indirect effect on the consistency constraint) reinforces the sectoral shift effect. The term $R$ captures any welfare effects arising from the reallocation of $V$ for a social planner with intrinsic redistributive preferences across sectors. As the following lemma shows, it disappears with relative welfare weights $\Psi(\theta, \varphi) = \tilde{\Psi}(F(\theta, \varphi))$:

**Lemma 3.** *With relative welfare weights, $\psi_E^\theta(w)/f_E^\theta(w) = \psi_E^\varphi(w)/f_E^\varphi(w)$ for all $w$, $E$.*

Since the planner only cares about redistribution across different wage earners, but not across individuals who earn the same wage in different sectors, there is no welfare effect from reallocating $V$ across sectors at given wages, and hence $R = 0$.

In addition to re-allocating across individuals at the same original wage, the $E$-induced decrease in wages shifts the population *on average* down the $V$ and $e$ schedules. The term $I$ comes from the incentive constraints and reflects the fact that the change in the aggregate distribution of $V$ and $e$ caused by this change in the wage distribution could not have been replicated via a change of taxes if the original wage distribution had been held constant. It is reminiscent of the effect that Stiglitz (1982) finds in a two-type model without occupational choice, but where wages are affected by each type's effort through a CRS production function. Then increasing the high skill type's effort has an extra incentive effect since it compresses the wage distribution. This is at work here as well, but complemented by the effects $S$, $C$ and $R$ from occupational choice and overlapping wage distributions in the two sectors, which are not captured by Stiglitz's two-type model.[13]

Setting $W'(E) = 0$ at any (interior) Pareto optimum yields the following relationship between $\xi$ and $t_{Pigou}$:

**Corollary 2.** *In any interior Pareto optimum,*

$$\xi = t_{Pigou} \left[ \frac{1 - \frac{1}{\lambda t_{Pigou}} \frac{\Delta\beta(E)}{\mu(E)} (I + R)}{1 + \frac{\Delta\beta(E)}{\mu(E)} (C + S)} \right]. \tag{14}$$

Hence, how $\xi$ (and thus the top marginal tax rate if the share of rent-seekers at the top is one) compares to the Pigouvian correction crucially depends on the sign of $\Delta\beta(E)$ and the four terms $I$, $R$, $C$ and $S$ from the outer problem. We consider a particularly simple case in the next subsection and the general case in subsection 3.6.

[13]Naito (1999) points out the role of sector-specific taxes in such models with pecuniary externalities, where manipulating wages can relax incentive constraints and therefore be desirable even if it introduces production inefficiencies.

### 3.5 A One-Sector Rent-Seeking Economy

We start by examining the benchmark case where all agents are rent-seekers and $f_E^\theta(w) = 0$ for all $w$ and $E$. This could be generated by concentrating all the skill density in the $\varphi$-dimension, resulting in the one-dimensional distribution $F(\varphi)$. Then obviously $I = R = C = S = 0$, so that (14) implies $\xi = t_{Pigou} = \beta^\varphi(E)$. Proposition 1 then implies the following marginal tax formulas in any Pareto optimum with Pareto weights $\Psi(\varphi)$:

**Proposition 2.** *Consider a one-sector rent-seeking economy. Then the marginal tax rate in any Pareto optimum without bunching is*

$$1 - T'(y(\varphi)) = (1 - t_{Pigou}) \left(1 + \frac{\eta(\varphi)}{\varphi f(\varphi)} \frac{1 + \varepsilon^u(\varphi)}{\varepsilon^c(\varphi)}\right)^{-1}$$

*with*

$$\eta(\varphi) = \int_\varphi^{\overline{\varphi}} \left(1 - \frac{\psi(x)}{f(x)} \frac{u_c(x)}{\lambda}\right) \exp\left(\int_\varphi^x \left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)}\right) \frac{dy(s)}{y(s)}\right) f(x) dx \tag{15}$$

*for all $\varphi \in \Phi$, where $\lambda$ is the multiplier on the resource constraint (4).*

The proposition shows that, in this case, the optimal correction $1 - t_{Pigou}$ is uniform, i.e. it does not depend on the skill level $\varphi$. This is an illustration of the "additivity property" discussed in the literature on optimal taxation in the presence of externalities (see Kopczuk, 2003, for the most general treatment).[14] The formula can be intuitively understood as a two-step correction: first tax all wages by $t_{Pigou}$ to correct the rent-seeking externality. Then apply standard optimal tax formulas, as in a Mirrlees model without externalities with the corrected wages $(1 - t_{Pigou})w$. In particular, the top marginal tax rate is just given by $T'(y(\overline{\varphi})) = t_{Pigou}$.

This intuition is somewhat misleading, in practice, since elasticities, incomes, and so forth, are endogenous. In the special case of quasilinear, isoelastic preferences (*viz* (8) with $\varepsilon(w) = \varepsilon$), however, the formula depends only on the *exogenous* distributions $F(\varphi)$ and $\Psi(\varphi)$ and the parameter $\varepsilon$. Proposition 2 therefore implies that the ratio of marginal keep shares $(1 - T'(y(\varphi)))/(1 - T'(y(\varphi')))$ for any $\varphi, \varphi'$ is independent of $t_{Pigou} = \beta_\varphi(E)$ in this case, and is the same as in a standard Mirrlees model. Hence, while the presence and magnitude of the rent-seeking externality affects the levels of optimal taxes via $1 - t_{Pigou}$, it does not affect the optimal progressivity of marginal tax rates in this case.

[14] See also Sandmo (1975), Sadka (1978), Bovenberg and van der Ploeg (1994), Pirttilä and Tuomala (1997) and Cremer, Gahvari, and Ladoux (1998). The additivity property holds whenever the externality is atmospheric and the externality generating good can be taxed directly. Since in our model the externality is generated by effort rather than a consumption good, we obtain a multiplicative decomposition of the tax system into its corrective and redistributive parts, but the underlying logic is the same as the additive structure found in the literature.

## 3.6 Top Marginal Tax Rates

As we show in the proof of the following proposition, these results change significantly when both sectors are present. Whether $\xi \gtreqless t_{Pigou}$ now crucially depends on the relative importance of the within- versus cross-sectoral externality, as captured by the sign of $\Delta\beta(E)$. This leads to the following result about the top marginal tax rate, which follows from Proposition 1 and Corollary 2.

**Proposition 3.** *Consider any regular Pareto optimum. Then $T'(y(\overline{w}_E)) \geq 0$ (strictly if $f_E^{\varphi}(\overline{w}_E) > 0$). If in addition (i) effort $e(w)$ is weakly increasing in $w$, (ii) marginal utility of consumption $u_c(c(w), e(w))$ is weakly decreasing in $w$, (iii) the share of rent seekers $f_E^{\varphi}(w)/f_E(w)$ is weakly increasing in $w$, (iv) the social welfare weight on traditional workers $\psi_E^{\theta}(w)/f_E^{\theta}(w)$ is weakly greater than that on rent-seekers $\psi_E^{\varphi}(w)/f_E^{\varphi}(w)$ at each $w$, and (v) the highest-wage workers are rent-seekers $f_E^{\varphi}(\overline{w}_E) = f_E(\overline{w}_E)$, then*

$$T'(y(\overline{w}_E)) \lesseqgtr t_{Pigou} \quad \text{if} \quad \Delta\beta(E) \gtreqless 0.$$

For instance, even if the top earners in the economy are *all* rent-seekers, the optimal top marginal tax rate is *less* than the full Pigouvian correction $t_{Pigou}$ if $\Delta\beta(E) > 0$. This contrasts with the result in Proposition 2 for a one sector economy, which implied that the top marginal tax rate was just equal to $t_{Pigou}$.

One might have expected a similar result to apply to the more general two sector model when all top earners are rent-seekers. Indeed, since rent-seeking imposes a negative externality and the government has a desire to redistribute from high to low earners, this seems like a clear case for high marginal tax rates on high earners. However, this intuition is at least partly undermined by the sectoral shift effect not present in the one sector model: lowering the marginal tax rate on the top earning rent-seekers below the Pigouvian correction increases total equivalent effort $E$. If $\Delta\beta(E) > 0$, the within-sector externality dominates the cross-sectoral externality, so wages in the rent-seeking sector fall, and by more than those in the traditional sector. As a consequence, some agents now find it profitable to exit the rent-seeking sector and become traditional workers. Since the traditional sector is socially more productive, this shift is welfare enhancing.

As discussed above, the increase in total rent-seeking effort $E$ has additional effects in a two-sector economy, which result from the fact that agents in both sectors must be treated the same conditional on the wage $w$, namely the effort re-allocation, incentive and welfare effects $C$, $I$ and $R$. The assumptions in Proposition 3 make sure that these effects go in the same direction as the sectoral shift effect, so that $C$, $I$ and $R$ are all positive. Note, however, that these are only sufficient conditions, so that $\xi < t_{Pigou}$ is possible even when they are violated for some wage levels. For instance, with relative Pareto weights and quasilinear preferences, $R = 0$ since the planner attaches the same welfare weight to individuals with the same wage in different sectors, and marginal utility of consumption is constant and equal to one, so that both conditions (ii) and (iv) can be dropped in this case. Moreover, of course we obtain $T'(y(\overline{w}_E)) < t_{Pigou}$ even if the share of rent-seekers is less than one at the top and $\Delta\beta(E) > 0$.

If $\Delta\beta(E) < 0$, all the effects reverse their sign. In this case, lowering the tax rate on the top earners would increase $E$ and lower wages in the traditional sector by more than in the rent-seeking sector, as the cross-sectoral externality is stronger than the within-sectoral one. It would therefore induce a wasteful sectoral short effect *into* the rent-seeking sector. As a result, the optimal top marginal tax rate *over-corrects* compared to the Pigouvian rate. This demonstrates that the optimal correction, accounting for the general equilibrium effects from occupational choice, deviates in a transparent way from the Pigouvian tax rate, outside of the knife-edge case $\Delta\beta(E) = 0$.

The reason why the additivity property fails in the two-sector case is that an income tax is not able to directly target the externality, which would require a sector specific tax on rent-seeking effort. Diamond (1973) considers another case where general equilibrium effects can lead to a deviation of the optimal correction from the Pigouvian rate. In his model, a congestion generating consumption good can be directly targeted, but only using a uniform linear tax even though individuals differ in their price and congestion sensitivity. Taxing the congestion good has a direct effect through the price sensitivity of demand, and an indirect, general equilibrium effect through the congestion sensitivity. He constructs an example where this leads to a low optimal corrective tax, similar in spirit to the above results.[15]

Our results readily extend to the case of a skill (and thus wage) distribution with unbounded support. For simplicity, consider the case with quasilinear and isoelastic preferences, so that (8) applies.[16] In addition, suppose that $\lim_{w\to\infty} f_E^{\varphi}(w)/f_E(w) = x$ with $x \in [0,1]$, so that the share of rent-seekers at the top of the income distribution converges to a constant. Moreover, assume that the wage distribution has Pareto tails, with parameter $\alpha$, i.e., that

$$\lim_{w\to\infty} \frac{1 - F_E(w)}{w f_E(w)} = \frac{1}{\alpha}.$$

[15]In contrast, our nonlinear income tax allows for a differential treatment of different skill types but excludes the direct taxation of the externality. Moreover, we have a distinct motivation for general equilibrium effects, and derive more transparent results about the departure from the Pigouvian rate. See also Micheletto (2008) for a deviation from the additivity property when the externality is not atmospheric.

[16]Similar results can be derived for the general case using the asymptotic methods in Saez (2001).

Finally, suppose that no welfare weight is put on the very top earners, so that $\psi_E(w)/f_E(w)$ goes to zero as $w$ grows. Then we can use equation (8) to derive the following asymptotic marginal tax rate for $w \to \infty$ (see Rothschild and Scheuer, 2011, for the details):

$$\lim_{w\to\infty} T'(y(w)) = \frac{\xi\alpha x + 1 + 1/\varepsilon}{\alpha + 1 + 1/\varepsilon}.$$

Moreover, Lemma 2 and Corollary 2 also go through, so that $0 < \xi \lesseqqgtr t_{Pigou}$ under the same conditions as in the bounded support case.

# 4 Conclusion

Our results can be attributed to the fact that income taxes are an imperfect instrument for correcting rent-seeking externalities. Directly taxing the externality-causing activity in the rent-seeking sector, were it possible, would reduce *both* its absolute desirability *and* its desirability relative to other activities. By contrast, an income tax directly affects only the *absolute* desirability of rent-seeking. The magnitude of the optimal correction via the income tax depends on the direction of the indirect (general equilibrium) effects of income taxes on the *relative* desirability of rent-seeking. When within-sector externalities dominate, these indirect effects are perverse: higher taxes on portions of the income distribution with high levels of rent-seeking activities raise the relative returns to rent-seeking and encourage entry into the rent-seeking sector. Consequently, the optimal externality correction lies strictly below the Pigouvian correction.

More generally, our results emphasize that the specific form of rent-seeking is crucial for optimal tax design: Are the rent-seekers' wages higher than their social marginal product because their effort depresses other rent-seekers's wages or the returns to productive activities? This is crucial in determining the effect of taxes on the relative return to rent-seeking and therefore on occupational choice. We do not attempt to provide quantitative evidence on this question, but view it as an important direction for future research.

# References

ACEMOGLU, D., AND T. VERDIER (1998): "Property Rights, Corruption and the Allocation of Talent: A General Equilibrium Approach," *Economic Journal*, 108, 1381–1403.

ARROW, K. (1973): "Social responsibility and economic efficiency," *Public Policy*, 21, 303–318.

BAUMOL, W. (1990): "Entrepreneurship: Productive, Unproductive, and Destructive," *Journal of Political Economy*, 98, 893–921.

BERTRAND, M., AND S. MULLAINATHAN (2001): "Are CEOs rewarded for luck?," *Quarterly Journal of Economics*, 116, 901–932.

BOVENBERG, L., AND L. GOULDER (1996): "Optimal Environmental Taxation in the Presence of Other Taxes: General-Equilibrium Analyses," *American Economic Review*, 86, 985–1000.

BOVENBERG, L., AND F. VAN DER PLOEG (1994): "Environmental policy, public finance, and the labor market in a second-best world," *Journal of Public Economics*, 55, 349–390.

CAHUC, P., AND E. CHALLE (2009): "Produce or Speculate? Asset Bubbles, Occupational Choice and Efficiency," *Mimeo, Ecole Polyechnique*.

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

CREMER, H., F. GAHVARI, AND N. LADOUX (1998): "Externalities and optimal income taxation," *Journal of Public Economics*, 70, 343–364.

DIAMOND, P. (1973): "Consumption Externalities and Imperfect Corrective Pricing," *Bell Journal of Economics and Management Science*, 4, 526–538.

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

DIXIT, A. (1985): "Tax policy in open economies," in *Handbook of Public Economics*, ed. by A. Auerbach, and M. Feldstein, vol. I, pp. 2475–2490. North-Holland, Amsterdam.

DIXIT, A. (1987): "Strategic behavior in contests," *American Economic Review*, 77, 891–898.

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

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.

KRUGMAN, P. (2009): "Rewarding Bad Actors," *New York Times*.

MICHELETTO, L. (2008): "Redistribution and Optimal Mixed Taxation in the Presence of Consumption Externalities," *Journal of Public Economics*, 92, 2262–2274.

MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.

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.

PHILIPPON, T. (2010): "Financiers versus Engineers: Should the Financial Sector be Taxed or Subsidized?," *American Economic Journal: Macroeconomics*, 2, 158–182.

PIRTTILÄ, J., AND M. TUOMALA (1997): "Income tax, commodity tax and environmental policy," *International Tax and Public Finance*, 4, 379–393.

ROCHET, J.-C., AND P. CHONÉ (1998): "Ironing, Sweeping and Multidimensional Screening," *Econometrica*, 66, 783–826.

ROTHSCHILD, C., AND F. SCHEUER (2011): "Optimal Taxation with Rent-Seeking," *NBER Working Paper 17035*.

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.

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

TULLOCK, G. (1967): "The welfare costs of tariffs, monopolies, and theft," *Western Economic Journal*, 5, 224–232.

——— (1980): "Efficient Rent-Seeking," in *Towards a Theory of the Rent-Seeking Society*, ed. by J. Buchanan, R. Tollison, and G. Tullock, pp. 97–112. Texas A&M University Press, College Station.

WERNING, I. (2000): "An Elementary Proof of Positive Optimal Marginal Tax Rates," *Mimeo, MIT*.

——— (2007): "Pareto-Efficient Income Taxation," *Mimeo, MIT*.

# A Proofs for Section 3

## A.1 Proof of Proposition 1

Putting multipliers $\lambda$ on (4), $\xi\lambda$ on (3) and $\hat{\eta}(w)\lambda$ on (2), the Lagrangian corresponding to (1)-(4) is, after integrating by parts (2),

$$\begin{aligned}\mathcal{L} \quad &= \textstyle\int_{\underline{w}_E}^{\overline{w}_E} V(w)\psi_E(w)dw - \int_{\underline{w}_E}^{\overline{w}_E} V(w)\hat{\eta}'(w)\lambda dw + \int_{\underline{w}_E}^{\overline{w}_E} u_e(c(V(w),e(w)),e(w))\frac{e(w)}{w}\hat{\eta}(w)\lambda dw \\ &+\xi\lambda\mu(E) - \xi\lambda \textstyle\int_{\underline{w}_E}^{\overline{w}_E} we(w)f_E^{\varphi}(w)dw + \lambda\int_{\underline{w}_E}^{\overline{w}_E} we(w)f_E(w)dw - \lambda\int_{\underline{w}_E}^{\overline{w}_E} c(V(w),e(w))f_E(w)dw. \qquad (16)\end{aligned}$$

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

$$\hat{\eta}'(w)\lambda = \psi_E(w) - \lambda f_E(w)\frac{1}{u_c(w)} + \hat{\eta}(w)\lambda\frac{u_{ec}(w)}{u_c(w)}\frac{e(w)}{w}. \tag{17}$$

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

$$\eta'(w) = \psi_E(w)\frac{u_c(w)}{\lambda} - f_E(w) + \eta(w)\frac{u_{cc}(w)c'(w) + u_{ce}(w)e'(w) + u_{ce}(w)e(w)/w}{u_c(w)}. \tag{18}$$

Using the first order condition corresponding to the incentive constraint (5),

$$u_c(w)c'(w) + u_e(w)e'(w) + u_e(w)\frac{e(w)}{w} = 0,$$

the fraction in (18) can be written as $-(\partial MRS(w)/\partial c)y'(w)/w$ where

$$MRS(w) \equiv -\frac{u_e(c(w),e(w))}{u_c(c(w),e(w))}$$

is the marginal rate of substitution between effort and consumption. Substituting in (18) and rearranging yields

$$-\frac{\partial MRS(w)}{\partial c}e(w)\frac{y'(w)}{y(w)}\eta(w) = f_E(w) - \psi_E(w)\frac{u_c(w)}{\lambda} + \eta'(w). \tag{19}$$

Integrating this ODE gives

$$\begin{aligned}\eta(w) &= \int_w^{\overline{w}_E}\left(f_E(w) - \psi_E(x)\frac{u_c(x)}{\lambda}\right)\exp\left(\int_w^x \frac{\partial MRS(s)}{\partial c}e(s)\frac{y'(s)}{y(s)}ds\right)dx \\ &= \int_w^{\overline{w}_E}\left(1 - \frac{\psi_E(x)}{f_E(x)}\frac{u_c(x)}{\lambda}\right)\exp\left(\int_w^x\left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)}\right)\frac{dy(s)}{y(s)}\right)f_E(x)dx, \qquad (20)\end{aligned}$$

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

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

$$\lambda w f_E(w)\left(1-\frac{MRS(w)}{w}\right)-\xi\lambda w f_E^{\varphi}(w)=-\hat{\eta}(w)\lambda\left[\frac{\left(-u_{ec}(w)u_e(w)/u_c(w)+u_{ee}(w)\right)e(w)}{w}+\frac{u_e(w)}{w}\right],$$

which after some algebra can be rewritten as

$$wf_E(w)\left(1-\frac{MRS(w)}{w}\right)-\xi w f_E^{\varphi}(w)=\eta(w)\left(\frac{\partial MRS(w)}{\partial e}\frac{e}{w}+\frac{MRS(w)}{w}\right). \tag{21}$$

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

$$1-\xi\frac{f_E^{\varphi}(w)}{f_E(w)}=\left(1-T'(y(w))\right)\left[1+\frac{\eta(w)}{wf_E(w)}\left(1+\frac{\partial MRS(w)}{\partial e}\frac{e}{MRS(w)}\right)\right]. \tag{22}$$

Simple algebra again shows that $1+\partial\log MRS(w)/\partial\log e = (1+\varepsilon^u(w))/\varepsilon^c(w)$, so that the result follows from (20) and (22).

## A.2 Proof of Lemma 1

Any Pareto optimum maximizes tax revenue subject to a set of minimum utility constraints for each type (with the Pareto weights as multipliers), the incentive constraints (2) and the consistency constraint (3). The last is equivalent to imposing the two inequality constraints $\mu(E)-\int_{\underline{w}_E}^{\overline{w}_E} we(w)f_E^{\varphi}(w)dw \geq 0$ and $\mu(E)-\int_{\underline{w}_E}^{\overline{w}_E} we(w)f_E^{\varphi}(w)dw \leq 0$. Establishing the lemma is equivalent to proving that the former binds (and hence the latter is slack) in any regular Pareto optimum. The proof proceeds by dropping the latter constraint and considering the optimum of the resulting relaxed problem.

Suppose, by way of contradiction, that $\xi = 0$ in this relaxed optimum. Standard arguments (e.g., Werning, 2000) imply $T'(y) \geq 0$ with regular Pareto weights in this case. Now consider a small decrease $\Delta E$ in $E$ holding fixed the tax schedule. This has no effect on the relaxed constraint (3). It has no effect on (2) since the tax schedule remains fixed. It at least weakly increases the wage, and hence the utility, of each type, relaxing the minimum utility constraints. (iv) Raises tax revenue, since $y(w)$ is non-decreasing in $w$ and $T'(y) \geq 0$. This contradicts the optimality of the allocation in the relaxed problem, showing that $\xi > 0$ in the relaxed problem. Hence $\mu(E)-\int_{\underline{w}_E}^{\overline{w}_E} we(w)f_E^{\varphi}(w)dw = 0$, the dropped constraint is satisfied, and the solution to the relaxed optimum—which has $\xi > 0$—coincides with the solution to the original problem.

## A.3 Proof of Lemma 2

Using (16),

$$\begin{aligned} W'(E) = &\int_{\underline{w}_E}^{\overline{w}_E} V(w)\frac{d\psi_E(w)}{dE}dw - \lambda\int_{\underline{w}_E}^{\overline{w}_E} c(V(w),e(w))\frac{df_E(w)}{dE}dw \\ &+\lambda(1-\xi)\int_{\underline{w}_E}^{\overline{w}_E} we(w)\frac{df_E^{\varphi}(w)}{dE}dw+\lambda\int_{\underline{w}_E}^{\overline{w}_E} we(w)\frac{df_E^{\theta}(w)}{dE}dw+\xi\lambda\mu'(E)+B_1 \end{aligned}$$

with

$$B_1 \equiv \quad \frac{d\overline{w}_E}{dE}\left[V(\overline{w}_E)\psi_E(\overline{w}_E) - \lambda c(V(\overline{w}_E), e(\overline{w}_E))f_E(\overline{w}_E) + \lambda\left(f_E(\overline{w}_E) - \xi f_E^{\varphi}(e(\overline{w}_E))\right)\overline{w}_E e(\overline{w}_E)\right]$$
$$-\frac{d\underline{w}_E}{dE}\left[V(\underline{w}_E)\psi_E(\underline{w}_E) - \lambda c(V(\underline{w}_E), e(\underline{w}_E))f_E(\underline{w}_E) + \lambda\left(f_E(\underline{w}_E) - \xi f_E^{\varphi}(e(\underline{w}_E))\right)\underline{w}_E e(\underline{w}_E)\right].$$

Integrating by parts the four integrals yields

$$W'(E) = B_1 + B_2 - \int_{\underline{w}_E}^{\overline{w}_E} V'(w)\frac{d\Psi_E(w)}{dE}dw + \lambda\int_{\underline{w}_E}^{\overline{w}_E}\left(\frac{V'(w)}{u_c(w)} + MRS(w)e'(w)\right)\frac{dF_E(w)}{dE}dw$$
$$-\lambda(1-\xi)\int_{\underline{w}_E}^{\overline{w}_E}(we'(w)+e(w))\frac{dF_E^{\varphi}(w)}{dE}dw - \lambda\int_{\underline{w}_E}^{\overline{w}_E}(we'(w)+e(w))\frac{dF_E^{\theta}(w)}{dE}dw + \xi\lambda\mu'(E) \quad (23)$$

with

$$B_2 = \left[V(w)\frac{d\Psi_E(w)}{dE} - \lambda c(V(w), e(w))\frac{dF_E(w)}{dE} + \lambda(1-\xi)we(w)\frac{dF_E^{\varphi}(w)}{dE} + \lambda we(w)\frac{dF_E^{\theta}(w)}{dE}\right]_{\underline{w}_E}^{\overline{w}_E}.$$

By the first order conditions (19) and (21) with respect to $V(w)$ and $e(w)$ from the inner problem, the terms

$$Q_1(E) \equiv \lambda\int_{\underline{w}_E}^{\overline{w}_E}\frac{e'(w)}{f_E(w)}\left[wf_E(w)\left(1-\frac{MRS(w)}{w}\right) - \xi wf_E^{\varphi}(w) - \eta(w)\left(\frac{\partial MRS(w)}{\partial e}\frac{e(w)}{w} + \frac{MRS(w)}{w}\right)\right]\frac{dF_E(w)}{dE}dw$$

and $\quad Q_2(E) \equiv \lambda\int_{\underline{w}_E}^{\overline{w}_E}\frac{V'(w)}{u_c(w)f_E(w)}\left[\psi_E(w)\frac{u_c(w)}{\lambda} - f_E(w) - \eta'(w) - \eta(w)\frac{\partial MRS(w)}{\partial c}e(w)\frac{y'(w)}{y(w)}\right]\frac{dF_E(w)}{dE}dw$

are both equal to zero. Adding $Q_1(E)$ and $Q_2(E)$ to (23), using (2) and re-arranging yields

$$W'(E) = B_1 + B_2 + \xi\lambda\mu'(E) + \int_{\underline{w}_E}^{\overline{w}_E} V'(w)\left(\frac{\psi_E(w)}{f_E(w)}\frac{dF_E(w)}{dE} - \frac{d\Psi_E(w)}{dE}\right)dw$$
$$-\lambda\int_{\underline{w}_E}^{\overline{w}_E} e(w)\frac{dF_E(w)}{dE}dw + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E}\left((e(w)+we'(w))\frac{dF_E^{\varphi}(w)}{dE} - we'(w)\frac{f_E^{\varphi}(w)}{f_E(w)}\frac{dF_E(w)}{dE}\right)dw$$
$$-\lambda\int_{\underline{w}_E}^{\overline{w}_E}\left(\frac{\eta(w)}{w}\frac{d\left[MRS(w)e(w)\right]}{dw} + \eta'(w)\frac{V'(w)}{u_c(w)}\right)\frac{1}{f_E(w)}\frac{dF_E(w)}{dE}dw. \quad (24)$$

Next, compute

$$\frac{dF_E(w)}{dE} = -\frac{d}{dE}\left(\frac{\mu(E)}{E}\right)\frac{E}{\mu(E)}wf_E^{\varphi}(w) - \frac{d\Gamma(E)}{dE}\frac{1}{\Gamma(E)}wf_E^{\theta}(w) = \frac{\beta^{\varphi}(E)}{E}wf_E^{\varphi}(w) + \frac{\beta^{\theta}(E)}{E}wf_E^{\theta}(w) \quad (25)$$

and analogously $d\Psi_E(w)/dE = \beta^{\varphi}(E)w\psi_E^{\varphi}(w)/E + \beta^{\theta}(E)w\psi_E^{\theta}(w)/E$. Moreover,

$$\frac{dF_E^{\varphi}(w)}{dE} = -\frac{d}{dE}\left(\frac{\mu(E)}{E}\right)\frac{E}{\mu(E)}wf_E^{\varphi}(w) + \frac{d}{dE}\left(\frac{\mu(E)}{E\Gamma(E)}\right)\int_{\underline{\varphi}}^{w\frac{E}{\mu(E)}}\varphi f\left(\varphi\frac{\mu(E)}{E\Gamma(E)}, \varphi\right)dw$$
$$\equiv \frac{\beta^{\varphi}(E)}{E}wf_E^{\varphi}(w) + K_E(w) \quad (26)$$

and analogously $dF_E^\theta(w)/dE = \beta^\theta(E) w f_E^\theta(w)/E - K_E(w)$. Then, the first integral in (24) is

$$\frac{\Delta\beta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} V'(w) w \frac{f_E^\theta(w) f_E^\varphi(w)}{f_E(w)}\left(\frac{\psi_E^\theta(w)}{f_E^\theta(w)} - \frac{\psi_E^\varphi(w)}{f_E^\varphi(w)}\right) dw = \frac{\Delta\beta(E)}{E} R \tag{27}$$

The terms with $e(w)$ on the second line of (24) can be written as

$$\begin{aligned}
&-\lambda(1-\xi)\frac{\beta^\varphi(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} w e(w) f_E^\varphi(w) dw - \lambda\frac{\beta^\theta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} w e(w) f^\theta(w) dw + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} e(w) K_E(w) dw \\
&\qquad = -\lambda\left(\beta^\varphi(E)\frac{\mu(E)}{E} + \beta^\theta(E)\frac{\Gamma(E)}{E} E_\theta\right) + \xi\lambda\beta^\varphi(E)\frac{\mu(E)}{E} + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} e(w) K_E(w) dw \\
&\qquad = -\lambda\frac{\mu(E)}{E} t_{Pigou} + \xi\lambda\beta^\varphi(E)\frac{\mu(E)}{E} + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} e(w) K_E(w) dw.
\end{aligned} \tag{28}$$

The terms with $we'(w)$ in (24) can be written as

$$\begin{aligned}
&\xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} we'(w)\left[\frac{\beta^\varphi(E)}{E} w f_E^\varphi(w) - \frac{f_E^\varphi(w)}{f_E(w)}\frac{w}{E}\left(\beta^\varphi(E) f_E^\varphi(w) + \beta^\theta(E) f_E^\theta(w)\right) + K_E(w)\right] dw \\
&\qquad = \xi\lambda\frac{\Delta\beta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} w^2 e'(w)\frac{f_E^\theta(w) f_E^\varphi(w)}{f_E(w)} dw + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} we'(w) K_E(w) dw \\
&\qquad = \xi\lambda\frac{\Delta\beta(E)}{E} C + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} we'(w) K_E(w) dw.
\end{aligned} \tag{29}$$

Combining the terms with $K_E(w)$ from (28) and (29) gives $\xi\lambda\int_{\underline{w}_E}^{\overline{w}_E}(we(w))' K_E(w) dw$, which can be integrated by parts to yield:

$$\begin{aligned}
&B_3 - \xi\lambda\frac{d}{dE}\left(\frac{\mu(E)}{E\Gamma(E)}\right)\frac{E}{\mu(E)}\int_{\underline{w}_E}^{\overline{w}_E} w^2 e(w) f\left(\frac{w}{\Gamma(E)}, \frac{wE}{\mu(E)}\right) dw \\
&\qquad = \xi\lambda\frac{\mu(E)}{\Gamma(E)}\frac{\Delta\beta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} w^2 e(w) f\left(\frac{w}{\Gamma(E)}, \frac{wE}{\mu(E)}\right) dw + B_3 = B_3 + \xi\lambda\frac{\Delta\beta(E)}{E} S.
\end{aligned} \tag{30}$$

with $B_3 = \xi\lambda\overline{w}_E e(\overline{w}_E) K_E(\overline{w}_E)$ since $K_E(\underline{w}_E) = 0$. Finally, use the incentive constraint (2), rewritten as $V'(w)/u_c(w) = MRS(w)e(w)$, to write the last line of (24) as

$$-\lambda\int_{\underline{w}_E}^{\overline{w}_E}\left(\eta(w) w\frac{d[V'(w)/u_c(w)]}{dw} + \eta'(w) w\frac{V'(w)}{u_c(w)} + \eta(w)\frac{V'(w)}{u_c(w)}\right)\frac{1}{w f_E(w)}\frac{dF_E(w)}{dE} dw$$

or, recognizing the sum of the bracketed terms as $d[\eta(w) w V'(w)/u_c(w)]/dw$, integrating by parts, and using the transversality condition $\eta(\underline{w}_E) = \eta(\overline{w}_E) = 0$ and (25),

$$\begin{aligned}
&\lambda\int_{\underline{w}_E}^{\overline{w}_E}\eta(w) w V'(w)\frac{d}{dw}\left(\frac{\beta^\theta(E)}{E}\frac{f_E^\theta(w)}{f_E(w)} + \frac{\beta^\varphi(E)}{E}\frac{f_E^\varphi(w)}{f_E(w)}\right) dw \\
&\qquad = \lambda\frac{\Delta\beta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E}\eta(w) w\frac{V'(w)}{u_c(w)}\frac{d}{dw}\left(\frac{f_E^\varphi(w)}{f_E(w)}\right) dw = \frac{\Delta\beta(E)}{E} I.
\end{aligned} \tag{31}$$

Define $\tilde{F}(w,E) \equiv F_E(w)$. Since $\tilde{F}(\overline{w}_E, E) \equiv 1$ for all $E$,

$$\frac{d\tilde{F}(\overline{w}_E,E)}{dE} = \frac{\partial\tilde{F}(\overline{w}_E,E)}{\partial E} + \frac{\partial\tilde{F}(\overline{w}_E,E)}{\partial w}\frac{d\overline{w}_E}{dE} = \frac{dF_E(\overline{w}_E)}{dE} + f_E(\overline{w}_E)\frac{d\overline{w}_E}{dE} = 0. \tag{32}$$

Together with an analogous expression at $\underline{w}_E$ and the fact that $K_E(\underline{w}_E) = 0$, this yields $B_1 + B_2 = -\xi\lambda\overline{w}_E e(\overline{w}_E)K_E(\overline{w}_E) = -B_3$. Using (27), (28), (29), (30) and (31) in (24) yields

$$W'(E) = -\lambda\frac{\mu(E)}{E}t_{Pigou} + \frac{\Delta\beta(E)}{E}(R+I) + \xi\lambda\left(\frac{\mu(E)}{E} + \frac{\Delta\beta(E)}{E}(C+S)\right), \tag{33}$$

where we have used $\mu'(E) + \beta^{\varphi}(E)\mu(E)/E = \mu(E)/E$.

## A.4 Proof of Lemma 3

Note first that

$$f_E^{\theta}(w) = \frac{1}{\Gamma(E)}\int_{\underline{\varphi}}^{w\frac{E}{\mu(E)}} f\left(\frac{w}{\Gamma(E)},\varphi\right)d\varphi = \frac{1}{\Gamma(E)}\frac{\partial F(w/\Gamma(E), wE/\mu(E))}{\partial\theta}$$

and analogously

$$f_E^{\varphi}(w) = \frac{E}{\mu(E)}\frac{\partial F(w/\Gamma(E), wE/\mu(E))}{\partial\varphi}.$$

With relative Pareto weights, $\Psi(\theta,\varphi) = \tilde{\Psi}(F(\theta,\varphi))$, so that

$$\psi_E^{\theta}(w) = \frac{1}{\Gamma(E)}\frac{\partial\Psi(w/\Gamma(E), wE/\mu(E))}{\partial\theta} = \frac{1}{\Gamma(E)}\tilde{\Psi}'\left(F\left(\frac{w}{\Gamma(E)},\frac{wE}{\mu(E)}\right)\right)\frac{\partial F(w/\Gamma(E), wE/\mu(E))}{\partial\theta}$$

and hence $\psi_E^{\theta}(w) = \tilde{\Psi}'(F_E(w))f_E^{\theta}(w)$. Analogously, $\psi_E^{\varphi}(w) = \tilde{\Psi}'(F_E(w))f_E^{\varphi}(w)$ with relative welfare weights. Hence, $\psi_E^{\theta}(w)/f_E^{\theta}(w) = \tilde{\Psi}'(F_E(w)) = \psi_E^{\varphi}(w)/f_E^{\varphi}(w)$ $\forall w, E$.

## A.5 Proof of Proposition 3

By Proposition 1, $T'(\overline{w}_E) = \xi f_E^{\varphi}(w)/f_E(w)$, and by Lemma 1, $\xi > 0$, establishing the first part of the proposition. Moreover, $\eta(w) \geq 0$ under the assumptions in the proposition. To see this, suppose (by way of contradiction) $\eta(w) < 0$ for some $w$. Since $\eta(\underline{w}_E) = \eta(\overline{w}_E) = 0$ by the transversality condition, this together with continuity of $\eta(w)$ implies that there exists some interval $[w_1, w_2]$ such that $w_1 < w_2$, $\eta(w_1) = \eta(w_2) = 0$ and $\eta(w) < 0$ for all $w \in (w_1, w_2)$. Then $\eta'(w_1) \leq 0$ and $\eta'(w_2) \geq 0$. Using (15), this implies

$$\frac{\psi_E(w_1)}{f_E(w_1)}\frac{u_c(w_1)}{\lambda} \leq \frac{\psi_E(w_2)}{f_E(w_2)}\frac{u_c(w_2)}{\lambda}.$$

However, $\psi_E(w)/f_E(w)$ is decreasing in $w$ with regular Pareto weights and $u_c(w)$ is also decreasing under condition (ii), yielding the desired contradiction. Hence, $I$ is non-negative under condition (iii). Conditions (i) and (iv) ensure that $C$ and $R$ are also non-negative, respectively, and $S > 0$. Hence, either the numerator or the denominator of (14) in Corollary 2 or both are positive. $\xi > 0$ implies that both are positive. Hence, $\xi \lesseqqgtr t_{Pigou} \Leftrightarrow \Delta\beta(E) \gtreqqless 0$.