CESifo Working Papers

www.cesifo.org/wp

Optimal Taxation with Rent-Seeking

Casey Rothschild
Florian Scheuer

CESifo Working Paper No. 5071
Category 1: Public Finance
November 2014

CESifo *Working Paper No. 5071*

Optimal Taxation with Rent-Seeking

Abstract

We develop a framework for optimal taxation when agents can earn their income both in traditional activities, where private and social products coincide, and in rent-seeking activities, where private returns exceed social returns either because they involve the capture of pre-existing rents or because they reduce the returns to traditional work. We characterize Pareto optimal non-linear taxes when the government does not observe the shares of an individual's income earned in each of the two activities. We show that the optimal externality correction typically deviates from the Pigouvian correction that would obtain if rent-seeking incomes could be perfectly targeted, even at income levels where all income is from rent-seeking. If rent-seeking externalities primarily affect other rent-seeking activity, then the optimal externality correction lies strictly below the Pigouvian correction. If the externalities fall mainly on the returns to traditional work, the optimal correction strictly exceeds it. We show that this deviation can be quantitatively important.

Keywords: rent-seeking, tax policy, multidimensional screening.

Casey Rothschild
Department of Economics
Wellesley College
USA - Wellesley, MA 02481
crothsch@wellesley.edu

Florian Scheuer
Department of Economics
Stanford University
USA - 94305-6072 Stanford CA
scheuer@stanford.edu

September 2014
We are especially grateful to Iván Werning for numerous valuable suggestions. We also thank Daron Acemoglu, Alan Auerbach, Marco Bassetto, Michael Boskin, Paco Buera, V. V. Chari, Peter Diamond, Emmanuel Farhi, Mikhail Golosov, Martin Hellwig, Caroline Hoxby, Patrick Kehoe, Phillip Levine, Guido Menzio, Chris Phelan, Thomas Piketty, James Poterba, Emmanuel Saez, Daniel Sichel, Matthew Weinzierl, David Wildasin and seminar participants at Alberta, Berkeley, MPI Bonn, Cologne, Copenhagen, Cornell, Frankfurt, Guelph, the Federal Reserve Bank of Minneapolis, Harvard, Konstanz, Mannheim, Munich, MIT, Stanford, Stockholm School of Economics, UMass-Amherst, Uppsala, Wellesley, Wharton, ETH Zurich, the SED Annual Meeting, NBER Summer Institute and Public Economics Program Meeting, Minnesota Workshop in Macroeconomic Theory, NTA Annual Conference on Taxation, the ASSA Meetings and the Young Macroeconomists' Jamboree (Duke) for helpful comments. Scheuer and Rothschild respectively thank Harvard University for hospitality and the Radcliffe Institute for Advanced Study at Harvard University for funding. All errors are our own.

1 Introduction

The financial crisis has exposed examples of highly compensated individuals whose apparent contributions to social output proved illusory. The view that some top incomes reflect rent-seeking—i.e., the pursuit of personal enrichment by extracting a slice of the existing economic pie rather than by increasing the size of that pie—has inspired calls for a more steeply progressive tax code (Piketty et al., 2014), and, motivated by similar concerns about rent-seeking in finance, various countries have proposed higher taxes on financial-sector bonus payments (Besley and Ghatak, 2013).

The argument behind such proposals is intuitively appealing. If part of the economic activity at high incomes is 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.[1] This would discourage such behavior while simultaneously raising revenues which could be used, for instance, to reduce taxes and encourage more productive effort at lower income levels. Moreover, if some sectors or professions are more prone to rent-seeking than others (Lockwood et al., 2014), sector-specific corrective taxes would be useful.

In this paper, we study the optimal design of such policies under the assumption of *imperfect targeting*. For example, lawyers produce many socially efficient services, upholding property rights and providing incentives to abide to useful rules. On the other hand, they may also engage in rent-seeking activities, some of which resemble zero-sum games. The friction that we account for here is that it can be very hard to tell which is which: the only way to find out can be a costly trial, a highly imperfect process. A similar point can be made of finance and many other sectors, as we discuss below. Hence, even sector- or profession-specific taxes (such as a bonus tax) are necessarily imperfectly targeted, as they apply to multiple different activities within such sectors that all come together in the same market and cannot be easily distinguished in any given transaction. At the extreme, an individual may engage both in productive and rent-seeking activities simultaneously, but only total income is observable when computing tax liabilities.

We adapt the Mirrlees (1971) framework to provide a formal foundation for studying the implications of such rent-seeking activities for optimal taxation under imperfect targeting. We use this framework to characterize optimal taxes in the presence of a broad class of rent-seeking externalities and to provide a precise formal comparison to the Pigouvian taxes that would obtain under perfect targeting. This allows us to provide conditions under which the above intuition about the progressivity of taxes is or is not

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

valid. We also simulate optimal tax schedules for an example economy calibrated to U.S. data, which allows us to quantitatively asses the impact of rent-seeking on tax policy.

We study an economy where individuals can pursue two types of activities: traditional work, where private and social returns to effort coincide, and rent-seeking, where private returns exceed the social returns to effort. We characterize the set of Pareto efficient income taxes, capturing the idea that the shares of an individual's income earned in rent-seeking and traditional work can be hard to disentangle. These taxes are characterized by a multiplicative correction to standard optimal tax formulas for economies without rent-seeking. Our main result is that this correction diverges systematically from the Pigouvian tax rate that would be optimal if rent-seeking income could be separately identified.

To illustrate this, suppose, for instance, that rent-seeking effort involves claiming credit for productive work done by other workers. Then rent-seeking imposes across-activity externalities (i.e., reduces the productivity of traditional effort) as well as within-activity externalities (due to crowding effects). Both externalities drive a wedge between the private returns to effort and its true underlying productivity, or social returns. There is, therefore, a potential role for corrective taxation. A natural, but typically incorrect guess for the optimal corrective component of the income tax would be the weighted Pigouvian tax rate, i.e. the average wedge between the private and social returns to effort at each income level.

To see why this guess is generally incorrect, consider raising marginal tax rates at income levels containing a high share of rent-seeking income. This directly reduces rent-seeking activity by discouraging effort at these incomes. But it also has *indirect* effects since a reduction in rent-seeking effort raises the returns to both types of effort. If the within-activity externalities are large relative to the across-activity ones, the returns to rent-seeking rise by more than the returns to productive effort. The tax change thus encourages a perverse shift of effort at all income levels *into* rent-seeking. This indirect effect partially offsets the direct corrective benefits of higher marginal taxes, and the optimal correction lies strictly below the Pigouvian rate.

When the across-activity rent-seeking externalities dominate the within-activity ones, on the other hand, a reduction in rent-seeking effort *lowers* the relative returns to rent-seeking, the activity shift effect reverses sign, and the optimal correction exceeds the Pigouvian tax rate. Only in the knife-edge case where the within and across-activity externalities exactly balance does the standard Pigouvian correction apply. Hence, the comparison depends on a transparent condition: does an additional unit of rent-seeking effort increase or decrease the relative returns to rent-seeking?

Solving for the optimal taxes in our framework is more challenging than in a standard Mirrlees (1971) optimal tax problem. The first reason is that wages are endogenous: for any given tax schedule, aggregate rent-seeking effort determines the returns to the two activities; and these returns determine individual and hence aggregate effort levels. The second challenge is that, naturally, individuals are characterized by a two-dimensional skill type: one skill for each activity. For these models, standard techniques typically do not apply (Rochet and Choné, 1998). Generalizing insights from Rothschild and Scheuer (2013), we address these challenges by observing that the realized wage distribution is well-defined conditional on any given aggregate rent-seeking effort (Lemma 2). Since taxes depend only on income, a standard single-crossing property allows us to treat the problem as a fixed point problem for aggregate rent-seeking effort nested within a Mirrleesian optimal 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 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 related problem without externalities.[2] In fact, if rent-seeking could be directly targeted, the optimal correction would be precisely the Pigouvian tax (Proposition 3).

Our main results, however, are driven by the observation that individuals in most industries engage in different activities, so taxes in practice can never be perfectly targeted: they cannot condition on the composition of individual incomes into income earned from rent-seeking and productive work. This restriction makes our analysis both theoretically interesting and practically relevant. In particular, Proposition 2 shows that the correction term in the optimal tax formula then diverges from the Pigouvian tax in manner that depends in a transparent way on the direction of the relative return effects.

We finally complement these analytical results with a quantitative exploration of *how large* this divergence can be in a calibrated version of our model using data from the 2013 Current Population Survey (CPS). In our baseline, the optimal correction amounts to only 53% of the Pigouvian correction, indicating that the general equilibrium effects we emphasize can be of first-order importance in practice, and of similar magnitude as the Pigouvian taxes themselves. We also perform comparative statics exercises, demonstrating the impact on tax policy of varying the importance of rent-seeking, e.g. reducing it to

[2] See Sandmo (1975), Sadka (1978), Cremer et al. (1998), Kopczuk (2003) and the discussion in section 4.6.

the level found in the model calibrated to CPS data from the 1970s. Similar to the recent quantitative work in Ales et al. (2014) for an economy without externalities, this allows us to decompose the change in optimal income taxes for the U.S. in the 1970s versus today into the component due to a change in rent-seeking and the component due to a change in the underlying skill distribution, holding redistributive motives fixed.

Related Literature. Our main result is related in spirit to Diamond (1973), although our motivation, framework, and instruments are quite distinct. Most importantly, Diamond considers linear commodity taxation in the Ramsey framework, whereas our analysis is in a Mirrleesian setting with non-linear income taxes, which combines redistributive and corrective motives for taxation. 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 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 two intensive margins corresponding to two income-earning activities. Moreover, in contrast to Diamond, 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.

A special case of our model obtains when rent-seeking income is earned through a crowdable search activity. Our analysis is therefore related to recent work by Golosov et al. (2013), who consider optimal taxation in labor markets with search frictions (see also Hungerbuhler et al., 2006). However, the former paper completely abstracts from skill-driven wage heterogeneity, in contrast to the Mirrleesian framework employed here. In addition, both papers consider search *for* employment rather than search *as* an income producing (but, through crowding, negative externality generating) activity.

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 (Kleven et al., 2009, Scheuer, 2013, 2014, Choné and Laroque, 2010, and Beaudry et al., 2009). These papers have different information structures than ours, however. The second dimension of heterogeneity enters preferences additively in the first three. In Choné and Laroque (2010), the second dimension is a taste for labor rather than a full second skill type as we employ here. In Beaudry et al. (2009), types have two distinct labor productivities, but one activity is a non-market activity, the returns from which are unobservable, whereas total income—but not its breakdown between the two activities—is observable in our model.[3] More closely related is Rothschild and Scheuer (2013), who use similar methods as developed here to characterize optimal taxation in a Roy (1951) model. That paper shares our structure of two-dimensional heterogeneity, but (as the other papers above) considers the special case where individuals always specialize in one type of activity and rules out wages that deviate from the social marginal product of effort and the resulting corrective motives for taxation, issues we focus on here.[4]

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 of them. Most of this literature (e.g. Baumol, 1990, Murphy et al., 1991, Acemoglu and Verdier, 1998, and Cahuc and Challe, 2012) does not consider optimal tax policy to correct these equilibrium outcomes. But there are important recent exceptions. Philippon (2010) considers an endogenous growth model with financiers, workers and entrepreneurs and analyzes the effect of linear, sector-specific taxes on growth. The recent studies by Piketty et al. (2014) and Lockwood et al. (2014) assume that rent-seeking reduces everyone else's income in a lump-sum fashion rather than the proportional reduction that we consider here. This rules out the relative return effects from effort that we emphasize, which arise when the externalities can affect different activities to varying degrees. Due to the absence of general equilibrium effects, the simple weighted Pigouvian correction is always optimal in their frameworks, which we use as a benchmark to compare our results to, both analytically and quantitatively. Furthermore, Lockwood et al. abstract from redistributive motives and Piketty et al. only consider top marginal tax rates, both in contrast to our analysis.

This paper proceeds as follows. In Section 2, we begin with a stylized example that generates a particularly stark version of our main result. Section 3 then introduces our general modeling framework and discusses applications and implementation issues. In Section 4, we analyze this model and provide our main results. Finally, Section 5 provides a quantification of these results for a calibrated version of our model, and Section 6 concludes. Most proofs as well as details on the data and calibration appear in a technical appendix.

[3]Beaudry et al. also assume that, unlike here, effort in the market activity is observable.

[4]Our setting with two unobservable margins of effort also relates to the literature on multi-tasking (Holmström and Milgrom, 1991 and Baker, 1992), although our Mirrleesian framework with private ex ante heterogeneity and externalities is quite distinct. In Rothschild and Scheuer (2014), we show how our method to solve the two-dimensional screening problem here can be extended to $N \geq 2$ dimensions.

2 A Simple Example

By way of illustration, we begin with a simple yet stark example to illustrate the key idea underlying our results. Consider an economy in which individuals pursue two types of activities. In the productive activity, output is proportional to (skill-weighted) aggregate effort. In the other activity, workers compete for a fixed rent $\bar{\mu} > 1$. There is an equal measure of two types of individuals, each with preferences $u(c,e) = c - e^\gamma/\gamma$ over consumption and effort. Type 1 agents have a skill $\theta_1 = 1$ for the productive activity and $\varphi_1 = 1$ for rent-seeking. In contrast, type 2 workers are unable to work in the productive activity ($\theta_2 = 0$) but have ability $\varphi_2 \equiv \varphi_R > 1$ for rent-seeking. Individuals of type i face returns to rent-seeking effort equal to $\varphi_i\bar{\mu}/E$, $i = 1,2$, where $E \equiv \lambda e_1 + \varphi_R e_2$ is the (skill-weighted) aggregate rent-seeking effort in the economy and λ is the (endogenous) fraction of type 1 workers who work in the rent-seeking activity. These rent-seeking returns correspond to a situation in which each unit of skill-weighted effort in that activity claims an equal share of the total rent $\bar{\mu}$.

We will show that the optimal nonlinear income tax can involve *zero* marginal tax rates for both types in this example economy, even though type 2 can only perform an activity that is socially completely wasteful. To wit, there are three possibilities for type 1's activity choice: If $E > \overline{\mu}$, the rent-seeking activity is relatively crowded, so type 1 individuals all prefer to do productive work ($\lambda = 0$); if $E < \overline{\mu}$, rent-seeking returns are higher than $\theta_1 = 1$, so type 1 workers strictly prefer to do rent-seeking ($\lambda = 1$); and finally an interior allocation with $E = \bar{\mu}$ where they are indifferent and some fraction $\lambda \in [0,1]$ of them does rent-seeking. We focus on this third case and derive parameter conditions below under which the utilitarian optimum indeed corresponds to such an interior allocation.

If $E = \overline{\mu}$, we can solve for the equilibrium share λ of type 1 workers doing rent-seeking

$$\overline{\mu} = E = \lambda e_1 + \varphi_R e_2 \quad \Rightarrow \quad \lambda = \frac{\overline{\mu} - \varphi_R e_2}{e_1}$$

and substituting yields a total income in the economy of $(1-\lambda)e_1 + \overline{\mu} = e_1 + \varphi_R e_2$. In other words, the wages of the two types are $w_1 = 1$ and $w_2 = \varphi_R$ since $E = \overline{\mu}$. As a result, utilitarian social welfare is simply

$$W = e_1 + \varphi_R e_2 - e_1^\gamma/\gamma + e_2^\gamma/\gamma. \tag{1}$$

Now suppose we have access to a nonlinear income tax, which allows us to control efforts e_i, $i = 1,2$, but not directly type 1's activity choices. Maximizing (1) w.r.t. e_i yields $e_i^{\gamma-1} =$

w_i, $i = 1, 2$, implying *zero* marginal tax rates for both types.[5] These efforts are consistent with an interior λ whenever $\varphi_R^{\gamma/(\gamma-1)} \in [\overline{\mu} - 1, \overline{\mu}]$. For sufficiently large γ, this interior allocation yields higher social welfare than the corner allocations with $E \neq \overline{\mu}$.[6] A fortiori, the optimal non-linear income tax for this economy has zero marginal taxes at both types' income levels.

In this example, the high income rent-seekers are clearly identifiable (a tax that hits only them can be levied, namely a tax on type 2's higher income $\varphi_R e_2$) and they produce no output. The targeted, Pigouvian tax on their effort would therefore be 100%. Yet it is optimal not to tax them under the nonlinear income tax. In other words, taxing the income of an identifiable subset of rent-seekers is not a good substitute for a direct tax on rent-seeking. To see why not, consider imposing a small tax on the type 2 individuals, reducing their effort by δ. This decrease in total rent-seeking effort E raises the returns to rent-seeking $\overline{\mu}/E$. Productive workers then shift into rent-seeking (λ increases) until $E = \overline{\mu}$ is restored. The net effect is an aggregate 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 activity, they help to keep type 1 workers out of rent-seeking and thereby sustain their productivity in the productive activity.

We call this indirect productivity the *activity 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 activity choice for optimal taxation in the presence of rent-seeking.[7] In the next two sections, we show that these effects are extremely robust and extend to a framework with general preferences and social welfare, with a continuum of types, with an arbitrary two-dimensional distribution of skills for the traditional and rent-seeking activities, wherein agents can work in both activities simultaneously, and wherein rent-seeking effort may impose externalities both on other rent-seekers—as in this example—and across-activity externalities on productive work. In Section 5, we also show that they can be of first-order importance quantitatively in more realistically calibrated versions of our model.

[5]Clearly, this zero-tax allocation is also incentive compatible, and therefore is the utilitarian optimum among the $E = \overline{\mu}$-allocations.

[6]To see this, note that social welfare in the optimum with interior λ is $W^* = \left(1 + \varphi_R^{\gamma/(\gamma-1)}\right)(\gamma - 1)/\gamma$. The highest welfare among the allocations with $\lambda = 1$ is $\overline{\mu}$, obtained in the allocation with 100% taxation and $E = 0$. Clearly, $W^* > \overline{\mu}$ for high enough γ since $\varphi_R^{\gamma/(\gamma-1)} > \overline{\mu} - 1$ as assumed above. The highest welfare with $\lambda = 0$ involves $E = \overline{\mu}$ and $e_1 = 1, e_2 = \overline{\mu}/\varphi^R$. It can be checked that the welfare from this allocation is always less than W^*.

[7]The same results obtain if type 1 individuals can only do productive work whereas type 2 individuals can pursue both activities.

3 The General Model

3.1 Setup

As in the example, individuals can pursue two activities: Traditional work, where private and social marginal products of effort coincide, and a rent-seeking activity, where the private marginal product exceeds the social marginal product. This model can apply to an entire economy or alternatively a particular sector, profession or industry within which both activities take place, as discussed in the introduction. There is a unit-measure continuum of individuals who choose how much effort to provide in each activity. Individuals are endowed with a two-dimensional skill vector $(\theta, \varphi) \in \Theta \times \Phi$, $\Theta = [\underline{\theta}, \overline{\theta}]$, $\Phi = [\underline{\varphi}, \overline{\varphi}]$, where θ captures an individual's skill in the traditional activity (the "Θ-activity"), and φ captures her skill in the rent-seeking activity (the "Φ-activity"). Skills are distributed with a continuous cdf $F : \Theta \times \Phi \to [0, 1]$ and associated continuous pdf $f(\theta, \varphi)$.

Preferences are characterized by a continuously differentiable and concave utility function over consumption c and effort in each activity, e_θ and e_φ, given by $U(c, e_\theta, e_\varphi) = u(c, m(e_\theta, e_\varphi)) \equiv u(c, e)$. We assume $u_c > 0$, $u_e < 0$, and that the effort aggregator $m(e_\theta, e_\varphi)$ is increasing in both arguments, continuously differentiable, quasiconvex and linear homogeneous.[8] A special case would obtain for $m(e_\theta, e_\varphi) = e_\theta + e_\varphi$, in which case individuals always specialize in the activity that yields them the highest returns, as in the Roy model considered in Rothschild and Scheuer (2013). Let the consumption and activity-specific efforts of an individual of type (θ, φ) be denoted by $c(\theta, \varphi)$, $e_\theta(\theta, \varphi)$, and $e_\varphi(\theta, \varphi)$. We respectively denote the total individual effort and utility of type (θ, φ) by $e(\theta, \varphi) \equiv m(e_\theta(\theta, \varphi), e_\varphi(\theta, \varphi))$ and $V(\theta, \varphi) \equiv u(c(\theta, \varphi), e(\theta, \varphi))$.

Aggregate output (income) is given by

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

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

are the aggregate effective (i.e., skill-weighted) efforts in the traditional and rent-seeking activities, respectively, and Y_θ and Y_φ are the aggregate incomes earned in the Θ- and Φ-activities as (continuously differentiable) functions of both aggregate efforts. The income of an individual of type (θ, φ) in each activity is $y_\theta(\theta, \varphi)$ and $y_\varphi(\theta, \varphi)$. Total individual in-

[8] Note that, since u is left general, this allows for preferences $\tilde{u}(c, \tilde{m}(e_\theta, e_\varphi))$ where $\tilde{m}$ is homothetic but not linear homogeneous: then there exist transformations u and m of $\tilde{u}$ and $\tilde{m}$ such that $\tilde{u}(c, \tilde{m}(e_\theta, e_\varphi)) = u(c, m(e_\theta, e_\varphi))$ for all (c, e_θ, e_φ) with linear homogeneous m. An example is $\hat{u}(c) - h_\theta(e_\theta) - h_\varphi(e_\varphi)$ with $h_\theta(.)$ and $h_\varphi(.)$ homogeneous of the same degree.

come is denoted by $y(\theta,\varphi) \equiv y_\theta(\theta,\varphi) + y_\varphi(\theta,\varphi)$, with $Y(E_\theta, E_\varphi) = \int_{\Theta\times\Phi} y(\theta,\varphi) dF(\theta,\varphi)$, and analogously for Y_θ and Y_φ.

To capture the distinction between the traditional and rent-seeking activities in terms of their social versus private returns, we model technology in this two-activity model as a fully general setting in which (i) each unit of effective effort in a given activity has the same private return, (ii) effort in the Θ-activity imposes no externalities, so its private return equals its social marginal product, and (iii) effort in the Φ-activity imposes at least weakly negative externalities on both activities. The following lemma formalizes these three properties and derives their implications for aggregate technology.

Lemma 1. *Suppose that there exist some $r_\theta(E_\theta, E_\varphi)$ and $r_\varphi(E_\theta, E_\varphi)$ such that*
(i) $y_\theta(\theta,\varphi) = r_\theta(E_\theta, E_\varphi)\theta e_\theta(\theta,\varphi)$ and $y_\varphi(\theta,\varphi) = r_\varphi(E_\theta, E_\varphi)\varphi e_\varphi(\theta,\varphi)$ $\forall\,(\theta,\varphi) \in \Theta\times\Phi$ (same private return to each unit of effective effort within a given activity),
(ii) $r_\theta(E_\theta, E_\varphi) = \partial Y_\theta(E_\theta, E_\varphi)/\partial E_\theta = \partial Y(E_\theta, E_\varphi)/\partial E_\theta$ $\forall E_\theta, E_\varphi$ (no externalities from the traditional activity), and
(iii) $r_\varphi(E_\theta, E_\varphi) \geq \partial Y_\varphi(E_\theta, E_\varphi)/\partial E_\varphi \geq \partial Y(E_\theta, E_\varphi)/\partial E_\varphi$ $\forall E_\theta, E_\varphi$ (weakly negative externalities from rent-seeking). Then

1. *$Y_\theta(E_\theta, E_\varphi) = \Gamma(E_\varphi)E_\theta$ for some $\Gamma(E_\varphi)$ with $\Gamma'(E_\varphi) \leq 0\ \forall\, E_\varphi$,*
2. *$Y_\varphi(E_\theta, E_\varphi) = \mu(E_\varphi)$ for some $\mu(E_\varphi)$ such that $\mu(E_\varphi)/E_\varphi \geq \mu'(E_\varphi)\ \forall\, E_\varphi$, and*
3. *$r_\theta(E_\theta, E_\varphi) = \Gamma(E_\varphi)$ and $r_\varphi(E_\theta, E_\varphi) = \mu(E_\varphi)/E_\varphi$.*

Proof. By (i), $Y_\theta(E_\theta, E_\varphi) = \int_{\Theta\times\Phi} y_\theta(\theta,\varphi)dF(\theta,\varphi) = \int_{\Theta\times\Phi} r_\theta(E_\theta, E_\varphi)\theta e_\theta(\theta,\varphi)dF(\theta,\varphi) = r_\theta(E_\theta, E_\varphi)E_\theta$ and, analogously, $Y_\varphi(E_\theta, E_\varphi) = r_\varphi(E_\theta, E_\varphi)E_\varphi$. Hence, $\partial Y_\theta(E_\theta, E_\varphi)/\partial E_\theta = r_\theta(E_\theta, E_\varphi) + E_\theta \partial r_\theta(E_\theta, E_\varphi)/\partial E_\theta$, and, by (ii), r_θ must be independent of E_θ. Since $\partial Y(E_\theta, E_\varphi)/\partial E_\theta = r_\theta(E_\theta, E_\varphi) + E_\varphi \partial r_\varphi(E_\theta, E_\varphi)/\partial E_\theta$, (ii) also implies that r_φ is independent of E_θ. We can thus write $Y_\theta(E_\theta, E_\varphi) = r_\theta(E_\varphi)E_\theta \equiv \Gamma(E_\varphi)E_\theta$ for some Γ, and $Y_\varphi(E_\theta, E_\varphi) = r_\varphi(E_\varphi)E_\varphi \equiv \mu(E_\varphi)$ for some μ. Finally, by (iii), $r_\varphi(E_\varphi) = \mu(E_\varphi)/E_\varphi \geq \mu'(E_\varphi) \geq \mu'(E_\varphi) + \Gamma'(E_\varphi)E_\theta\ \forall E_\theta,\ E_\varphi$, which requires $\mu(E_\varphi)/E_\varphi \geq \mu'(E_\varphi)$ and $\Gamma'(E_\varphi) \leq 0$ for all E_φ. □

Since private returns to effective effort only depend on E_φ by result 3 in Lemma 1, we simplify notation by dropping the subscript and letting $E \equiv E_\varphi$ henceforth. Lemma 1 establishes that in any two-activity model with properties (i) to (iii), traditional and rent-seeking outputs can respectively be written as $\Gamma(E)E_\theta$ and $\mu(E)$. Moreover, private returns to effective effort must be $r_\theta(E) = \Gamma(E)$ in the traditional activity and $r_\varphi(E) = \mu(E)/E$ in the rent-seeking activity, with $\Gamma'(E) \leq 0$ and $\mu(E)/E \geq \mu'(E)$.[9]

Returns in the traditional activity reflect the social marginal product of effort in that activity, while private returns exceed the social marginal product of effort in the rent-seeking activity for two reasons: First, rent-seeking effort imposes a negative *cross-activity*

[9]This is guaranteed, e.g., if μ is increasing and concave with $\mu(0) \geq 0$, although it is not necessary.

externality on the traditional activity by reducing output in that activity when $\Gamma'(E) < 0$. Second, rent-seeking effort imposes a negative *within-activity* externality when $r_\varphi(E) = \mu(E)/E > \mu'(E)$ and the private returns exceed the within-activity social marginal product. One extreme case would arise if $\Gamma(E) \equiv 1$ and $\mu(E) \equiv E$, so that the rent-seeking externality disappears. On the other hand, pure rent-seeking occurs when $\Gamma(E) \equiv 1$ and $\mu(E) = \overline{\mu}$, so there is a fixed rent to be captured in the rent-seeking activity and any effort there is socially unproductive "skimming," as in the example in Section 2.

It is worth emphasizing that this technology does not require firms or employers. We can assume that each worker is self-employed and reaps the return to his effort directly. However, the returns to both productive and rent-seeking effort are determined in general equilibrium, by the supply of effort of all other workers.

While this technology thus generally captures a situation in which one activity exhibits negative externalities and the other does not, property (i) of the Lemma implies that the rent-seeking externality works through individual returns to effective effort. It therefore rules out the uniform *absolute* reduction in other individuals' incomes due to rent-seeking (considered, for instance, in Piketty et al., 2014, and Lockwood et al., 2014), which is independent of effort. Indeed, the interesting relative return effects on activity choice that we explore in the following arise precisely because we allow rent-seeking to have differential effects on the returns to different types of effort.

3.2 Applications

Our general framework is flexible enough to capture a wide range of rent-seeking activities. For instance, in many applications, labor effort effectively takes the form of search, such as search for arbitrage opportunities in financial markets, where traders compete to exploit a potentially limited set of profitable trades. As a trading strategy becomes more crowded, its equilibrium return falls, leading to negative search externalities (as e.g. in Axelson and Bond, 2012). Similarly, portions of individual incomes are often earned through tournaments (Lazear and Rosen, 1979), i.e., fixed-sum games with winner-takes-all compensation, e.g. in activities related to the arts, entertainment, law, or consulting. The next example illustrates our connection to these standard rent-seeking applications formally.

Example 1 (Contests). *Consider N rent-seekers competing for a rent of value μ. As in Tullock (1980), the probability p_i that player $i \in \{1, ..., N\}$ wins the contest is increasing in her own* *effective effort relative to everyone else's effective effort:*

$$p^i(\varphi^i e^i_\varphi, \varphi^{-i} e^{-i}_\varphi) = \varphi^i e^i_\varphi \Big/ \sum_{j=1}^{N} \varphi^j e^j_\varphi.$$

Player i's expected payoff (and thus long-run income if these contests occur repeatedly) is therefore $\varphi^i e^i_\varphi \mu / E$ *with* $E \equiv \sum_{j=1}^{N} \varphi^j e^j_\varphi$. *Whenever* $\varphi^i e^i_\varphi / E$ *is small, the private return to effective rent-seeking effort is given by* μ / E, *as in our general model, and exceeds the social marginal return to effective effort, given by* $\mu'(E)$ *(zero if* $\mu(E) = \overline{\mu}$*).*

Relatedly, our model can capture activities that take the form of races, where individuals compete to be the first to discover new ideas and innovations in R&D and academic research, market opportunities or consulting strategies (see e.g. Arrow, 1962, for business stealing effects and Dixit, 1987, for innovation races). This framework also describes socially wasteful but privately profitable financial speculation (Hirshleifer, 1971, Arrow, 1973) when traders have an incentive to be the first to incorporate a piece of information into market prices, even if the social value of this acceleration is small, leading to excessive investments in accelerating the pace of adjustments. This is again demonstrated formally in the following example.

Example 2 (Races)**.** *Suppose individuals race to discover a rent with value* $M(t)$ *at time t (with* $M' < 0$ *if early discovery is valuable), with the winner capturing the entire benefit. If individual discovery hazards* $\lambda \varphi^i e^i_\varphi$ *are proportional to skill-weighted efforts* $\varphi^i e^i_\varphi$, *then, defining aggregate effort* $E = \sum_i \varphi^i e^i_\varphi$, *the time to discovery is* $p(t|E) = \lambda E \exp(-\lambda E t)$ *and the expected payoff to an individual rent-seeker i is*

$$\frac{\varphi^i e^i_\varphi}{E} \int_0^\infty M(t) p(t|E) dt \equiv \varphi^i e^i_\varphi \frac{\mu(E)}{E} \quad \text{with} \quad \mu(E) \equiv \int_0^\infty M(t) p(t|E) dt,$$

as in our general model. Conditions that ensure $\mu(E)/E \geq \mu'(E)$ *are easy to provide.*

Of course, many jobs involve a mixture of both rent-seeking and traditional, productive activities. For instance, Glode and Lowery (2012) consider a model where financial sector workers engage in both (zero-sum) speculative trading and surplus creation (e.g. from market making) and argue that profits from both activities are interlinked. In such cases, rent-seeking may also have negative impacts on the productivity of traditional activities, as captured by $\Gamma'(E) < 0$ in our general model and illustrated in the next example, which elaborates on the illustrative example from the introduction.

Example 3 (Red Tape)**.** *Consider an organization wherein individuals compete for rents* μ *(as in Example 1) and provide traditional work. Agents skim rents by routing paperwork through* *their desk or inserting themselves more in the way of decision and production processes. The total income of individual i is given by* $\varphi^i e^i_\varphi \mu / E + \Gamma(E)\theta^i e^i_\theta$, *where* $\Gamma'(E) < 0$ *captures the negative effect on traditional output when individuals compete harder for rents.*

Rent-seeking is also a natural feature of labor markets where pay is determined by offers from competing employers, such as for academic faculty. Academics can put time and effort both into doing actual research and into competing for outside offers (which is costly and time-consuming both for the academic and potential recruiters). Both types of effort increase pay, but salary raises due to outside offers may not necessarily correspond to an increase in research productivity, as captured by the contest compensation structure in Example 1.[10] Spending time and resources on outside offers may also get in the way of actual research, leading to cross-activity externalities as in Example 3.

While Example 3 describes negative production externalities from rent-seeking, our framework can also capture situations where some of the rent-seeking incomes are at the expense of productive incomes. For instance, Besley and Ghatak (2013) argue that bailouts in the financial sector come at the expense of productive workers, leading to wasteful (risk-taking) effort among financial workers with negative cross-sectoral externalities. Piketty et al. (2014) emphasize that executive officers may set their compensation through bargaining, thereby claiming a larger share of the resources in the company and leaving less for workers. In Biais et al. (2011), fast traders impose externalities on slow traders through adverse selection from their information advantage. In team production settings, individuals spend time and effort both to produce profits and to get credit for those profits (and to make sure to get compensated for the profits they get credit for).

As these examples make clear, our assumption that rent-seeking incomes are hard to target separately not only makes our analysis theoretically interesting but is also most relevant in practice. In particular, even sector- or profession-specific taxes e.g. on the financial sector, such as the bonus taxes mentioned in the introduction, affect individuals who earn income both from productive work and rent-seeking. Our model then characterizes the optimal design of a non-linear tax for any given such sector or occupation. As discussed in the introduction, enforcing even more targeted instruments, such as taxes only on particular activities within financial markets or law services, would require the government to collect detailed information about the kind of tasks that individuals do and therefore be difficult and costly, especially if individuals can rather easily relabel their type of activity.[11]

[10]The same framework applies when researchers spend part of their time on competing for grants or prizes through races or contests with winner-takes-all compensation, as in Examples 1 and 2, and other parts of their time on producing actual research.

[11]Our qualitative results would extend to the case where incomes from different activities can be dis-

In an alternative interpretation, the model can be applied to an entire economy with a general income tax and where activities themselves correspond to sectors or occupations. The activity shift effect could then be interpreted as a shift of effort across such sectors. Our assumption of imperfect targeting within sectors would then translate into the assumption that profession-specific tax instrument are not available (as in Rothschild and Scheuer, 2013, and Lockwood et al., 2014). This is also the interpretation we pursue in the quantitative exercise in Section 5. One could imagine various practical reasons for such a restriction. Differential taxation may create additional distortions if individuals can relabel their occupations or shift income from one sector to another. It may encourage special interest lobbying for preferential tax treatment of particular occupations. Finally, it may raise concerns about horizontal equity and about empowering the government to make the determination of how socially productive workers in different professions or sectors really are. A general income tax also corresponds to the tax systems in most countries, so the second interpretation would connect naturally to the recent policy debate.[12]

3.3 Implementation

We start by characterizing a general, direct implementation, where individuals announce their privately known type (θ, φ) and then get assigned $c(\theta, \varphi)$, $y(\theta, \varphi)$, and the fraction of income earned through the Θ-activity, given by $q(\theta, \varphi) \equiv y_\theta(\theta, \varphi)/y(\theta, \varphi) = \Gamma(E)\theta e_\theta(\theta, \varphi)/y(\theta, \varphi)$. Income y and consumption c are observable but an individual's skill type (θ, φ) and, due to imperfect targeting, their activity-specific efforts e_θ or e_φ (and q) are not. The resulting incentive constraints that guarantee truth-telling of the agents are:

$$u\left(c(\theta,\varphi), m\left(\frac{q(\theta,\varphi)y(\theta,\varphi)}{\theta\Gamma(E)}, \frac{(1-q(\theta,\varphi))y(\theta,\varphi)}{\varphi\mu(E)/E}\right)\right)$$
$$\geq \max_{p\in[0,1]}\left\{u\left(c(\theta',\varphi'), m\left(\frac{py(\theta',\varphi')}{\theta\Gamma(E)}, \frac{(1-p)y(\theta',\varphi')}{\varphi\mu(E)/E}\right)\right)\right\} \ \forall(\theta,\varphi),(\theta',\varphi'), \qquad (2)$$

since type (θ, φ) can imitate another type (θ', φ') by earning (θ', φ')'s income using a continuum of combinations of efforts (and thus income shares $(p, 1-p)$) in the Θ- and

tinguished imperfectly, e.g. because of an imperfect signal about an individual's income composition or a "fuzzy frontier" between different types of income, which allows for income shifting. This issue has received much attention and empirical support in the context of labor versus capital income (e.g. Saez et al., 2012) and would presumably apply even more to income from traditional versus rent-seeking effort.

[12] We take the rent-seeking opportunities and hence externalities in a given sector as given. Presumably, the government can also affect rents through regulation, which would affect the form of the externalities through $\Gamma(E)$ and $\mu(E)$ directly. However, as long as some rent-seeking opportunities remain after regulation, our analysis of optimal taxes remains relevant, taking the degree of regulation as given.

Φ-activities.

The following lemma shows that incentive compatibility implies that each type (θ, φ) has a well-defined wage $w \equiv y/e$ and activity-specific income share q, which both depend on aggregate rent-seeking effort E but are otherwise independent of the allocation.

Lemma 2. *In any incentive compatible allocation* $\{c(\theta,\varphi), y(\theta,\varphi), q(\theta,\varphi), E\}$,

$$w(\theta,\varphi) \equiv \frac{y(\theta,\varphi)}{e(\theta,\varphi)} = \max_{p\in[0,1]} m\left(\frac{p}{\theta\Gamma(E)}, \frac{1-p}{\varphi\mu(E)/E}\right)^{-1} \tag{3}$$

and $q(\theta,\varphi)$ *is a corresponding* arg max.

Proof. By the linear homogeneity of m, the "own type" incentive constraints (for $(\theta,\varphi) = (\theta',\varphi')$) imply

$$q(\theta,\varphi) \quad \in \quad \underset{p\in[0,1]}{\operatorname{argmin}}\ m\left(\frac{py(\theta,\varphi)}{\theta\Gamma(E)}, \frac{(1-p)y(\theta,\varphi)}{\varphi\mu(E)/E}\right) = \underset{p\in[0,1]}{\operatorname{argmin}}\ m\left(\frac{p}{\theta\Gamma(E)}, \frac{1-p}{\varphi\mu(E)/E}\right).$$

Equation (3) then follows immediately from the definitions $w(\theta,\varphi) \equiv y(\theta,\varphi)/e(\theta,\varphi)$ and

$$e(\theta,\varphi) \equiv m(e_\theta(\theta,\varphi), e_\varphi(\theta,\varphi)) = y(\theta,\varphi) m\left(\frac{q(\theta,\varphi)}{\theta\Gamma(E)}, \frac{1-q(\theta,\varphi)}{\varphi\mu(E)/E}\right).$$ □

By Lemma 2, fixing E pins down each type's wage $w(\theta,\varphi)$; we write $w_E(\theta,\varphi)$ to make this E-dependency explicit henceforth. Moreover, $q(\theta,\varphi)$ is chosen to minimize the overall effort $m(e_\theta, e_\varphi)$ subject to achieving a given amount of income. Hence, by (3) and linear homogeneity of m,

$$w_E(\theta,\varphi) = \max_{p\in[0,1]} ym\left(\frac{py}{\theta\Gamma(E)}, \frac{(1-p)y}{\varphi\mu(E)/E}\right)^{-1} = \max_{e_\theta,e_\varphi} \frac{y}{m(e_\theta,e_\varphi)} \text{ s.t. } \theta\Gamma(E)e_\theta + \varphi\frac{\mu(E)}{E}e_\varphi = y \tag{4}$$

for any y. When m is strictly quasiconvex, $q(\theta,\varphi)$, which depends only on the skill ratio $\alpha \equiv \theta/\varphi$, is also uniquely determined by E for each α. With weakly quasiconvex m, $q(\theta/\varphi)$ is uniquely determined outside of a countable set of α-values, where it is interval-valued because the individual is indifferent between various effort combinations to achieve a given income. By Lemma 2, the correspondence $q(\theta/\varphi)$ depends on, and is non-decreasing in, the *relative* returns $x_E(\alpha) \equiv \alpha\Gamma(E)/(\mu(E)/E)$, and we can define the *functions* $q_E(\alpha) \equiv Q(x_E(\alpha))$ by taking q to be equal to the midpoints of the intervals on the countable set of degenerate α's. Viewed as a *distribution*, Q has a well-defined derivative denoted by Q'.[13] Note that the constraint in (4) can be rearranged to

$$\frac{e_\theta}{e_\varphi} = \frac{\varphi\mu(E)/E}{\theta\Gamma(E)}\left(\frac{y}{e_\varphi\varphi\mu(E)/E} - 1\right) = \frac{1}{x_E(\alpha)}\frac{q}{1-q}, \tag{5}$$

[13] In particular, $Q(x)$ is ordinarily differentiable except at some countable number of jump discontinuities; the latter adds a series of Dirac δ-functions to Q'. See appendix A.2. Recall that $q \in \{0,1\}$ for almost all individuals if m is linear, but otherwise will typically take interior values.

so that $Q(x) = xr(x)/(1 + xr(x))$ with $r \equiv e_\theta/e_\varphi$, which only depends on and is increasing in x by linear homogeneity and quasiconvexity of m. For later use, it will also be helpful to define $\tilde{Q}_{x_0}(x) \equiv x_0 r(x)/(1 + x_0 r(x))$ with $\tilde{Q}'(x) \equiv \tilde{Q}'_{x_0}(x)\big|_{x_0=x}$, i.e. the change in the traditional income share Q that is due to the change in the activity-specific effort ratio r in response to a change in relative returns x, but holding x fixed otherwise.

All individuals with the same wage w have the same preferences over (c, y)-bundles given by $u(c, y/w)$. As is standard, we assume the single crossing property, i.e., that the marginal rate of substitution between y and c, $-u_e(c, y/w)/\ (wu_c(c, y/w))$, is decreasing in w. Under this assumption, it is straightforward to show that any incentive compatible allocation can be implemented with a non-linear income tax $T(y)$ and that we can trace out the Pareto frontier by restricting attention to allocations $\{c(w), y(w), E\}$ that pool all same-wage individuals at the same (c, y)-bundle (see Lemma 1 in Rothschild and Scheuer, 2013).

4 Optimal Non-linear Income Taxation

4.1 Definitions and Preliminaries

We use general cumulative Pareto weights $\Psi(\theta, \varphi)$ in (θ, φ)-space 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. The observation that makes this problem tractable is that, by Lemma 2, fixing E determines the wage $w_E(\theta, \varphi)$ and the traditional activity income share $q_E(\alpha)$ for each type (θ, φ). For any given E, it will be useful in the following to compute the cdf over (w, α)-pairs from

$$G_E(w, \alpha) \equiv \int_{\{(\theta,\varphi)|w_E(\theta,\varphi)\leq w, \theta/\varphi\leq\alpha\}} dF(\theta, \varphi)$$

with the corresponding density $g_E(w, \alpha)$. We denote the support of the wage distribution for any E by $[\underline{w}_E, \overline{w}_E]$, where $\underline{w}_E = w_E(\underline{\theta}, \underline{\varphi})$ and $\overline{w}_E = w_E(\overline{\theta}, \overline{\varphi})$. This allows us to obtain the wage distribution for any given E simply as the marginal distribution

$$F_E(w) \equiv \int_{\{(\theta,\varphi)|w_E(\theta,\varphi)\leq w\}} dF(\theta, \varphi) = \int_{\underline{w}_E}^{w} \int_{\alpha=0}^{\infty} dG_E(z, \alpha)$$

with the corresponding density $f_E(w) = \int_{\alpha=0}^{\infty} dG_E(w, \alpha)$ as well as the activity-specific densities $f_E^\theta(w) \equiv \int_{\alpha=0}^{\infty} q_E(\alpha) dG_E(w, \alpha)$ and $f_E^\varphi(w) \equiv \int_{\alpha=0}^{\infty} (1 - q_E(\alpha)) dG_E(w, \alpha)$. Hence, these densities can be interpreted as an average value of q (respectively $1 - q$) for all individuals at wage w, and $f_E(w) = f_E^\theta(w) + f_E^\varphi(w)$ for all $w \in [\underline{w}_E, \overline{w}_E]$.[14] Finally, for any given E, we can derive Pareto weights over wages $\Psi_E(w)$, as well as their density and decomposition across activities $\psi_E(w) = \psi_E^\theta(w) + \psi_E^\varphi(w)$, completely analogously from $\Psi(\theta, \varphi)$. We are particularly interested in the regular case in which the planner assigns greater weight to low-wage individuals, i.e., where $\psi_E(w)/f_E(w)$ is non-increasing in w for any E.[15]

Any incentive compatible allocation $\{c(w), y(w), E\}$ implies total effort and utility $e(w) \equiv y(w)/w$ and $V(w) \equiv u(c(w), e(w))$ as well as the activity-specific efforts $e_\theta(\theta, \varphi) = q_E(\theta/\varphi) y(w_E(\theta, \varphi))/\theta\Gamma(E)$ and (analogously) $e_\varphi(\theta, \varphi)$. We denote the uncompensated and compensated wage elasticities of total effort e by $\varepsilon^u(w)$ and $\varepsilon^c(w)$, respectively.

4.2 A Decomposition and Pareto Optimality

Generalizing the analysis of the extensive-margin Roy model in Rothschild and Scheuer, 2013, we can decompose the problem of finding Pareto optimal allocations into two steps. The first step involves finding the optimal level of aggregate rent-seeking effort E. We call this the "outer" problem. The second (which we call the "inner" problem) involves finding the optimal resource-feasible and incentive-compatible allocation for a given level of E. This inner problem is an almost standard Mirrlees problem; the only difference is that the induced level of aggregate effective rent-seeking effort has to be consistent with the level of E that we are fixing for the inner problem. For some given Pareto weights $\Psi(\theta, \varphi)$ (and hence induced 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. c):

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

subject to

$$V'(w) + u_e(c(V(w), e(w)), e(w)) \frac{e(w)}{w} = 0 \;\; \forall w \in [\underline{w}_E, \overline{w}_E] \tag{7}$$

$$\mu(E) - \int_{\underline{w}_E}^{\overline{w}_E} w e(w) f_E^\varphi(w) dw = 0 \tag{8}$$

[14] In the special case with $m(e_\theta, e_\varphi) = e_\theta + e_\varphi$, (3) immediately implies $q_E(\alpha) \in \{0, 1\}$ almost everywhere and $w_E(\theta, \varphi) = \max\{\theta\Gamma(E), \varphi\mu(E)/E\}$. Then, as in Rothschild and Scheuer (2011), $f_E^\varphi(w)/f_E(w)$ can be interpreted as the share of rent-seekers at w (whereas here it is more generally the rent-seeking income share at wage w).

[15] For example, consider the case of *relative* Pareto weights where $\Psi(\theta, \varphi) = \tilde{\Psi}(F(\theta, \varphi))$ for some increasing function $\tilde{\Psi} : [0, 1] \to [0, 1]$. Then these Pareto weights are regular whenever $\tilde{\Psi}$ is weakly concave.

$$\int_{\underline{w}_E}^{\overline{w}_E} we(w) f_E(w) dw - \int_{\underline{w}_E}^{\overline{w}_E} c(V(w), e(w)) f_E(w) dw \geq 0. \tag{9}$$

We employ the standard Mirrleesian approach of optimizing directly over allocations, i.e., over effort $e(w)$ and consumption or, equivalently, utility $V(w)$ profiles. The social planner maximizes a weighted average of individual utilities $V(w)$ subject to three constraints. (9) is a standard resource constraint and (8) ensures that aggregate effective effort in the rent-seeking activity indeed sums up to E (or, equivalently, rent-seeking incomes sum to $\mu(E)$). Finally, the allocation $V(w), e(w)$ needs to be incentive compatible, i.e.,

$$V(w) \equiv u(c(w), e(w)) = \max_{w'} u\left(c(w'), \frac{e(w')w'}{w}\right). \tag{10}$$

It is a well-known result that under single-crossing, the global incentive constraints (10) are equivalent to the local incentive constraints (7) and the monotonicity constraint that income $y(w)$ must be non-decreasing in w.[16] We follow the standard approach of dropping the monotonicity constraint, which can easily be checked ex post (as we do for the numerical simulations in Section 5). If the solution to problem (6) to (9) does not satisfy it, optimal bunching would need to be considered. Accounting for bunching is conceptually straightforward and does not substantively effect our analysis, so, for simplicity, we abstract from bunching henceforth.

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 $\max_E W(E)$. It is straightforward to show that a solution to the inner problem exists for any E (see Rothschild and Scheuer, 2014, for details) and that, at any E for which individuals work in both activities, $W(E)$ is continuous, so that the outer problem has a solution over any compact set of Es.[17]

4.3 Marginal Tax Rate Formulas from the Inner Problem

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

Proposition 1. *The marginal tax rate in any Pareto optimum is such that*

$$1 - T'(y(w)) = \left(1 - \xi \frac{f_E^{\varphi}(w)}{f_E(w)}\right) \left(1 + \frac{\eta(w)}{w f_E(w)} \frac{1 + \varepsilon^u(w)}{\varepsilon^c(w)}\right)^{-1} \qquad \textit{with} \tag{11}$$

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

[17] Compactness would be ensured, for instance, by a standard Inada condition $u_e(c, e) \to -\infty$ as $e \uparrow \bar{e}$ for some $\bar{e} < \infty$.

$$\eta(w) = \int_w^{\overline{w}_E} \left(1 - \frac{\psi_E(x)}{f_E(x)} \frac{u_c(x)}{\lambda}\right) \exp\left(\int_w^x \left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)}\right) \frac{dy(s)}{y(s)}\right) f_E(x) dx \tag{12}$$

for all $w \in [\underline{w}_E, \overline{w}_E]$, where λ is the multiplier on the resource constraint (9) and $\lambda\xi$ the multiplier on the consistency constraint (8).

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 ξ is the (normalized) Lagrange multiplier on constraint (8) and $f_E^{\varphi}(w)/f_E(w)$ is the fraction of incomes earned in the rent-seeking activity at wage level w. This optimal local correction, which makes agents internalize the rent-seeking externality, is thus proportional to the relative importance of rent-seeking at w and the shadow cost of the consistency constraint (8). As usual, the term $\eta(w)$ captures the redistributive motives of the government and income effects from the terms in the exponential function. A particularly transparent formula can be obtained from (11) with quasilinear preferences $u(c,e) = c - h(e)$, 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 $T'(y(w)) \geq 0$ at all income levels under regular Pareto weights, and the marginal tax rate is increasing in the degree to which $\Psi_E(w)$ shifts weight to low-wage individuals compared to $F_E(w)$.

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

4.4 Optimal Rent-Seeking Effort from the Outer Problem

Our main goal here is to compare ξ to the Pigouvian tax t_{Pigou}, defined by

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

i.e., as the tax that aligns the private and social returns to rent-seeking effort. We show in section 4.6 below that t_{Pigou} can be interpreted as the optimal corrective tax if, in addition to levying an optimal nonlinear income tax, the government could directly tax rent-seeking income (see Proposition 3). The key question in the following will be how ξ—interpretable the optimal externality correction in our model with imperfect targeting—*differs* from this targeted instrument benchmark t_{Pigou}. For this purpose, it is useful to denote the elasticities of the returns $r_\theta(E)$, $r_\varphi(E)$ in the two activities w.r.t. E by

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

Then t_{Pigou} can be expressed as an output weighted sum of the corrections for the within- and the cross-activity externalities from rent-seeking:

$$t_{Pigou} = \beta_\varphi(E) + \frac{Y_\theta}{Y_\varphi}\beta_\theta(E) > 0.$$

Let $\Delta\beta(E) \equiv \beta_\varphi(E) - \beta_\theta(E)$ denote the relative importance of the within- versus across-activity externalities. Lemma 3 uses this to provide a simple decomposition of the welfare effect of marginal changes in E.

Lemma 3. *The welfare effect of a marginal change in aggregate rent-seeking effort E is*

$$W'(E) = \lambda\frac{\mu(E)}{E}\left(\xi - t_{Pigou}\right) + \frac{\Delta\beta(E)}{E}\left[I + R + \xi\lambda\left(C+S\right)\right], \tag{13}$$

where

$$I \equiv \lambda\int_{\underline{w}_E}^{\overline{w}_E}\eta(w)w\frac{V'(w)}{u_c(w)}\frac{d}{dw}\left(\frac{f_E^\varphi(w)}{f_E(w)}\right)dw, \tag{14}$$

$$R \equiv \int_{\underline{w}_E}^{\overline{w}_E}V'(w)w\frac{f_E^\theta(w)f_E^\varphi(w)}{f_E(w)}\left(\frac{\psi_E^\theta(w)}{f_E^\theta(w)} - \frac{\psi_E^\varphi(w)}{f_E^\varphi(w)}\right)dw, \tag{15}$$

$$C \equiv \int_{\underline{w}_E}^{\overline{w}_E}w^2e'(w)Var_E(q|w)f_E(w)dw \tag{16}$$

and

$$S \equiv \int_{\underline{w}_E}^{\overline{w}_E}\int_{\alpha=0}^{\infty}y(w)\tilde{Q}'(x_E(\alpha))x_E(\alpha)dG_E(w,\alpha) \geq 0 \tag{17}$$

with $Var_E(q|w) = \int_0^\infty q_E(\alpha)^2g_E(\alpha|w)d\alpha - \left(\int_0^\infty q_E(\alpha)g_E(\alpha|w)d\alpha\right)^2$ *and* $g_E(\alpha|w) = \frac{g_E(w,\alpha)}{f_E(w)}$.

The terms $\lambda\xi\mu(E)/E$ and $-\lambda t_{Pigou}\mu(E)/E$ in (13) respectively capture the direct effect of a change in E on constraint (8) and the effect of changing sectoral returns on (9), holding fixed the bundles $e_\theta(\theta,\varphi)$, $e_\varphi(\theta,\varphi)$, and $V(\theta,\varphi)$ (and hence $c(\theta,\varphi)$) for all types (θ,φ).[18] In fact, when $\Delta\beta(E) = 0$, E has no effect on relative returns $x_E(\alpha)$. So changing E while

[18]To see the former, multiply (8) by $1/r_\varphi(E) = E/\mu(E)$ to re-write it as $E - \int_{\Theta\times\Phi}\varphi e_\varphi(\theta,\varphi)dF(\theta,\varphi) = 0$ and note that the second term is independent of E when $e_\varphi(\theta,\varphi)$ is held constant. To see the latter, differentiate with respect to E the total income in the economy $r_\theta(E)E_\theta + r_\varphi(E)\int_{\Theta\times\Phi}\varphi e_\varphi(\theta,\varphi)dF(\theta,\varphi)$ to get

$$\frac{r_\theta'(E)E}{r_\theta(E)}\frac{r_\theta(E)E_\theta}{E} + \frac{r_\varphi'(E)E}{r_\varphi(E)}r_\varphi(E) = -\beta_\theta(E)\frac{Y_\theta}{E} - \beta_\varphi(E)r_\varphi(E) = -\frac{\mu(E)}{E}t_{Pigou}.$$

holding the effort-consumption bundle for each type (θ, φ) fixed is compatible with the incentive constraints (2). By an envelope argument, then, $W'(E) = \lambda \left(\xi - t_{Pigou}\right) \mu(E)/E$, consistent with (13).

When $\Delta\beta(E) \neq 0$, the change in relative returns drives a wedge between ξ and t_{Pigou} since holding allocations fixed is no longer incentive compatible, and there are additional welfare effects from a change in E captured by the four effects in (14) to (17). In discussing them, we focus on the case $\Delta\beta(E) > 0$, so that an increase in E increases the relative return to traditional work $x_E(\alpha)$. The opposite case is analogous with reversed signs.

First, a change in E causes an *activity shift*. An increase in E (and thus $x_E(\alpha)$) leads a type $\alpha = \theta/\varphi$-individual to increase her optimal effort ratio $e_\theta/e_\varphi = r(x)$ and hence the traditional income share $Q(x) = xr(x)/(1 + xr(x))$. The change in Q due to the change in r (holding x itself constant) is $\tilde{Q}'(x_E(\alpha))dx_E(\alpha)/dE = \tilde{Q}'(x_E(\alpha))x_E(\alpha)\Delta\beta(E)/E$, whence the term S in (13). As we detail in Appendix A.2, there are two components to this change: a continuous one, which arises as individuals continuously increase e_θ/e_φ (and hence q), and a discrete one, which arises when the effort aggregator $m(e_\theta, e_\varphi)$ has linear segments and an incremental change in E leads some indifferent individuals to jump discretely to a higher q (as emphasized in Rothschild and Scheuer, 2011, where just indifferent individuals switch the activity they specialize in). The (distributional) derivative $\tilde{Q}'$ subsumes both of these effects.

Second, a change in E has *different* effects on the wages of distinct individuals who originally earned the *same* wage w. In particular, note that (4) can equivalently be written as

$$w_E(\theta, \varphi) = \max_{e_\theta, e_\varphi} \frac{\theta\Gamma(E)e_\theta + \varphi\frac{\mu(E)}{E}e_\varphi}{m(e_\theta, e_\varphi)} \text{ s.t. } m(e_\theta, e_\varphi) = e. \tag{18}$$

Using the envelope theorem, the elasticity of wages with respect to E is therefore

$$-\frac{dw_E(\theta, \varphi)}{dE}\frac{E}{w_E(\theta, \varphi)} = q_E(\alpha)\beta_\theta(E) + (1 - q_E(\alpha))\beta_\varphi(E), \tag{19}$$

i.e., the income-share weighted average of the aggregate return elasticities. The *average* wage change for individuals originally pooled at w is thus $-\left(\bar{q}_w\beta_\theta + (1 - \bar{q}_w)\beta_\varphi\right) w/E$, where we write $\bar{q}_w$ as a shorthand for $\mathbb{E}_E(q|w)$. By (19) and when $\Delta\beta(E) > 0$, α-types with $q_E(\alpha) < (>)\, \bar{q}_w$ see their wages fall by more (less) than this average, however. The terms C, I, and R are the extra welfare effects that arise in reconciling this with the fundamental incentive constraints (2). Specifically, since the thought experiment of holding type-specific allocations fixed is infeasible when $\Delta\beta(E) \neq 0$, the formal proof we provide in Appendix A.2 (and discuss in more intuitive detail in Rothschild and Scheuer, 2011) is motivated by instead holding fixed the *wage*-specific allocations $e(w), V(w)$.

The term C arises because, by changing their wages differentially in the face of a fixed effort schedule $e(w)$, an increase in E effectively re-allocates effort across individuals with the same initial wage w but different activity-specific intensities q. In particular, the change in effort induced by E, relative to the wage-w average, is $e'(w)(\bar{q}_w - q)\Delta\beta(E)w/E$ for an individual with original wage w and traditional income share q. The change in rent-seeking income for this individual is thus $w^2e'(w)(1-q)(\bar{q}_w - q)\Delta\beta(E)/E$. Averaging over all q's yields $w^2e'(w)Var_E(q|w)\Delta\beta(E)/E$ and hence (after summing over all w's) the effect C on (8). In particular, for an increasing effort schedule $e(w)$, a rise in E results in a re-allocation of effort from low- to high-q individuals at any given w—and thus from the Φ- to the Θ-activity. Hence, C reinforces S if $e'(w) \geq 0$.

The terms R and I are parallel to those in Rothschild and Scheuer (2013). R arises from the analogous reallocation of utility $V(w)$ from low-q to high-q individuals at the same initial wage, as a rise in E decreases the latter's wage by relatively less when $\Delta\beta(E) > 0$. It obviously disappears with relative welfare weights $\Psi(\theta,\varphi) = \tilde{\Psi}(F(\theta,\varphi))$, since then $\psi_E^\theta(w)/f_E^\theta(w) = \psi_E^\varphi(w)/f_E^\varphi(w)$ for all w, E. Otherwise, it is welfare improving when the planner puts more weight on high-q individuals at each wage ($\psi_E^\theta(w)/f_E^\theta(w) > \psi_E^\varphi(w)/f_E^\varphi(w)$) and vice versa.

Finally, the term I is a generalized Stiglitz (1982) effect: if $\Delta\beta(E) > 0$ and the share of income earned through rent-seeking is locally increasing in w (i.e., $d(f_E^\varphi(w)/f_E(w))/dw > 0$), an increase in E leads to a local compression of the wage distribution, since returns in the high-wage activity fall and vice versa. This yields a welfare improving easing of the local incentive constraints (7) if they are binding downwards ($\eta(w) \geq 0$). I therefore vanishes if there are no redistributive motives (e.g. with quasilinear preferences and $\tilde{\Psi}(F) = F\ \forall F$), so that $\eta(w) = 0$ for all w.

With linear m, the model collapses to the special case in Rothschild and Scheuer (2011) with binary occupational choice, $q \in \{0,1\}$. Then the effects simplify as follows:

Corollary 1. *If $m(e_\theta, e_\varphi) = e_\theta + e_\varphi$, then I and R are as in Lemma 3 and*

$$C = \int_{\underline{w}_E}^{\overline{w}_E} w^2 e'(w) \frac{f_E^\theta(w) f_E^\varphi(w)}{f_E(w)} dw, \qquad S = \frac{E}{\mu(E)\Gamma(E)} \int_{\underline{w}_E}^{\overline{w}_E} w^2 e(w) f\left(\frac{w}{\Gamma(E)}, \frac{wE}{\mu(E)}\right) dw \geq 0.$$

C then just captures the re-allocation of effort from rent-seekers to traditional workers with the same initial w at wage levels where the activity-specific distributions overlap, and S measures the effort moved from Φ to Θ as originally indifferent individuals switch activities.

4.5 Marginal Tax Rate Results

Setting $W'(E) = 0$ and using (13) at any interior Pareto optimum yields the following relationship between ξ and t_{Pigou}:

$$\xi = t_{Pigou}\left(1 - \frac{1}{\lambda t_{Pigou}}\frac{\Delta\beta(E)}{\mu(E)}(I+R)\right) \Bigg/ \left(1 + \frac{\Delta\beta(E)}{\mu(E)}(C+S)\right). \tag{20}$$

In a one-activity model with only the rent-seeking activity available and $f_E^\theta(w) = 0$ for all w, we mechanically have $I = R = C = S = 0$ and therefore $\xi = t_{Pigou} = \beta_\varphi(E)$. The tax formula (11) then implies that the correction factor by which marginal keep shares are scaled down compared to the standard formula is *uniform* and given by $1 - t_{Pigou}$. This can be understood as a two-step correction as in Kopczuk (2003): first tax all wages by t_{Pigou} to correct the rent-seeking externality. Then apply the standard optimal tax formula, as in a Mirrlees model without externalities, with the corrected wages $(1 - t_{Pigou})w$. In particular, the top marginal tax rate is just $T'(y(\overline{w}_E)) = t_{Pigou}$.

In the general case where both activities take place, the optimal correction ξ deviates from t_{Pigou} due to the relative return effects (14) to (17) whenever $\Delta\beta(E) \neq 0$. Based on the discussion in the previous subsection and (20), the following proposition collects conditions that determine this comparison.

Proposition 2. *In any regular Pareto optimum, $\xi > 0$. 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-seeking incomes $f_E^\varphi(w)/f_E(w)$ is weakly increasing in w, and (iv) the welfare weights on high-q workers are weakly greater than those on low q-workers at each w, so that $\psi_E^\theta(w)/f_E^\theta(w) \geq \psi_E^\varphi(w)/f_E^\varphi(w)\ \forall w$, then*

$$\xi \lesseqqgtr t_{Pigou} \quad \text{if} \quad \Delta\beta(E) \gtreqqless 0.$$

Combined with the marginal tax rate formula in Proposition 1, this result has clear implications for Pareto optimal tax schedules. For instance, under the conditions in Proposition 2 and if all income at the top is earned through rent-seeking ($f_E^\varphi(\overline{w}_E) = f_E(\overline{w}_E)$), then $T'(y(\overline{w}_E)) = \xi \lesseqqgtr t_{Pigou}$ if $\Delta\beta(E) \gtreqqless 0$. Hence, if e.g. $\Delta\beta(E) > 0$, the top marginal tax rate is less than the Pigouvian correction t_{Pigou} even when all top earners are exclusively active in the rent-seeking activity.[19] At other income levels, the optimal correction ξ is still less than t_{Pigou} by Proposition 2, but of course gets combined with the redistributive

[19]Of course, a fortiori we obtain $0 \leq T'(y(\overline{w}_E)) = \xi f_E^\varphi(\overline{w}_E)/f_E(\overline{w}_E) < t_{Pigou}$ if the share of income from rent-seeking is less than one at the top and $\Delta\beta(E) > 0$.

components of the marginal tax rate according to (11).

The divergence of the optimal correction ξ from t_{Pigou} directly reflects the fact that the income tax is an imperfect tool for externality correction. It is an imperfect tool *even* in brackets of the income distribution where *all* income comes from rent-seeking. This is because, as the discussion above highlights, the effects of the externality E are non-uniform—i.e., there are *relative* return effects whenever $\Delta\beta(E) \neq 0$. When $\Delta\beta(E) > 0$, taxing rent-seeking intensive portions of the income distribution at a higher rate directly discourages effort at those income levels, lowering E and helping to correct the externality. Partly offsetting this is the fact that a lower E *raises* the relative returns to rent-seeking, encouraging a shift *into* this activity. This partial offset implies a smaller-than-Pigouvian optimal correction. If $\Delta\beta(E) < 0$, all the effects reverse their sign: higher taxes directly discourage the externality-causing activity, and, since this lowers the relative returns to rent-seeking, indirectly encourage effort-shifting away from rent-seeking. As a result, the optimal ξ exceeds t_{Pigou} and, for instance, the optimal top marginal tax rate *over-corrects* compared to the Pigouvian rate.

As discussed in subsection 4.4, apart from $S \geq 0$, there are three additional relative return effects C, I, and R. The assumptions in Proposition 2 are sufficient to ensure that they are also non-negative and thus reinforce S. Note, however, that these are only sufficient conditions, so that the comparison between ξ and t_{Pigou} can hold 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 but different q's, and marginal utility of consumption is constant and equal to one, so that both conditions (ii) and (iv) can be dropped.[20] Assumptions (i) and (iii) are easy to verify ex-post, as we do in Section 5.

Our results do not depend on a bounded skill distribution, but readily extend to the case of an unbounded support. For simplicity, consider quasilinear and isoelastic preferences.[21] In addition, suppose that $\lim_{w\to\infty} f_E^{\varphi}(w)/f_E(w) = x$ with $x \in [0,1]$, that $\chi = \lim_{w\to\infty} w f_E(w)/(1-F_E(w))$ exists, and that $\lim_{w\to\infty} \psi_E(w)/f_E(w) = 0$ so that the share of rent-seeking income at the top is well-defined, the wage distribution has a Pareto tail, and the social planner puts zero weight on the top earners. Then we can use equation (11) to derive the following asymptotic marginal tax rate for $w \to \infty$ (see Rothschild and

[20]The only role of condition (ii) in Proposition 2 is to make sure (together with regular welfare weights) that the incentive constraints bind downwards, i.e. $\eta(w) \geq 0$. All that matters for this is that the *overall* social marginal welfare weights $u_c(w)\psi_E(w)/f_E(w)$ are non-increasing in w.

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

Scheuer, 2011, for the details):

$$\lim_{w\to\infty} T'(y(w)) = \frac{\xi\chi x + 1 + 1/\varepsilon}{\chi + 1 + 1/\varepsilon}. \tag{21}$$

Moreover, Lemma 3 and Proposition 2 also go through, so that $0 < \xi$ and $\xi \lesseqqgtr t_{Pigou}$ under the same conditions as in the bounded support case.

4.6 Comparison to the Perfectly Targeted Tax Benchmark

We finally compare the preceding results to the hypothetical scenario under which, in addition to the nonlinear income tax $T(y)$, a linear tax t on rent-seeking income is available. To attack the resulting optimal tax problem, it is useful to decompose it into an inner and outer problem as follows: the inner problem takes t and E as given and is written in terms of after-t returns in the rent-seeking activity. In particular, we define allocations as before, except that now $y_\varphi(\theta,\varphi) \equiv (1-t)\varphi\frac{\mu(E)}{E}e_\varphi(\theta,\varphi)$, with $y(\theta,\varphi) \equiv y_\theta(\theta,\varphi) + y_\varphi(\theta,\varphi)$ and $q(\theta,\varphi) \equiv y_\theta(\theta,\varphi)/y(\theta,\varphi)$ following accordingly. Then by Lemma 2, incentive compatibility requires

$$w_{E,t}(\theta,\varphi) \equiv \max_{p\in[0,1]} m\left(\frac{p}{\theta\Gamma(E)}, \frac{1-p}{(1-t)\varphi\mu(E)/E}\right)^{-1},$$

$q_{E,t}(\theta/\varphi)$ is a corresponding arg max, and both are determined by t and E for each type (θ,φ). This leaves the objective (6) and incentive constraints (7) in the inner problem unchanged. The consistency and resource constraints become

$$\int_{\underline{w}_{E,t}}^{\overline{w}_{E,t}} we(w) f^{\varphi}_{E,t}(w)dw + t\mu(E) = \mu(E) \quad \text{and} \tag{22}$$

$$\int_{\underline{w}_{E,t}}^{\overline{w}_{E,t}} \left(we(w) - c(V(w),e(w))\right) f_{E,t}(w)dw + (t-s)\mu(E) + \frac{s}{1-t}\int_{\underline{w}_{E,t}}^{\overline{w}_{E,t}} we(w) f^{\varphi}_{E,t}(w)dw = 0, \tag{23}$$

with multipliers $\xi\lambda$ and λ as before, and where s is a free parameter that will be useful in what follows (the terms multiplied by s are identically 0 by (22), and a casual reader can safely set $s = 0$). The inner problem yields social welfare $W(E,t)$, which the outer problem maximizes over E and t. The next proposition summarizes the results for this scenario with targeted tax instruments.

Proposition 3. *Suppose a linear activity-specific tax t on rent-seeking is available in addition to the non-linear income tax $T(y)$. Then, at any Pareto optimum:*
(i) $(1-t)\xi = t_{Pigou} - (t-s)$,

(ii) the top marginal tax rate on rent-seeking income is $t + (t_{Pigou} - t) f^{\varphi}_{E,t}(\overline{w}_{E,t}) / f_{E,t}(\overline{w}_{E,t})$, *which reduces to* t_{Pigou} *if all top earners specialize in rent-seeking, and*
(iii) if there are no redistributive motives ($\eta(w) = 0 \, \forall w$, $\Psi(\theta, \varphi) = F(\theta, \varphi) \, \forall (\theta, \varphi)$)*, then* $t = t_{Pigou}$ *and no other distorting taxes are imposed* ($T'(y) \equiv 0$ *for all* y)*.*

Proposition 3 formally motivates our choice of t_{Pigou} as the benchmark corrective tax based on three insights: First, as shown in part (ii), t_{Pigou} is the top marginal tax rate when all top-wage workers only do rent-seeking *and* we can target rent-seeking directly. Second, part (iii) implies that, in the absence of redistributive motives (e.g. with quasilinear preferences and $\Psi = F$), the optimal tax on the rent-seeking component of income is given by t_{Pigou}.

Third, although the optimal linear rent-seeking tax t is *not* generally equal to t_{Pigou} when there are redistributive motives, t_{Pigou} can still be interpreted its *corrective component*. Specifically, part (i) implies that, as in Kopczuk (2003), we can solve for an optimum in two steps. First, set a baseline corrective tax τ. Second, solve for the remaining linear rent-seeking tax $s = t - \tau$ and the non-linear income tax T. If we choose $\tau = t_{Pigou}$, part (i) yields $\xi = 0$, so the optimization problem in the second step is equivalent to a problem with *no* externalities (but with pre-s returns $(1 - t_{Pigou}) \varphi \mu(E) / E$ in the rent-seeking activity). Hence, t_{Pigou} is the Pigouvian corrective tax t^p as defined by Kopczuk (2003).

Proposition 3 also makes transparent that our key results are due to the fact that it is impossible to target the rent-seeking externality directly. If a tax on rent-seeking were available, the optimal correction would simply be given by the standard Pigouvian correction t_{Pigou}. In contrast, our analysis and the results in Proposition 2 focus on the more realistic case where such a direct instrument is not available, since only total income y—but not its composition q—is verifiable, and overall rent-seeking effort in a given profession or industry can only indirectly be influenced through the design of the (profession-specific) nonlinear income tax. The interesting general equilibrium effects from endogenous wages and activity choices play a crucial role for shaping optimal tax policy only under this restricted instrumentarium.

5 Numerical Illustration

In this section, we provide optimal policy simulations for a simple version of our model calibrated to the U.S. in order to quantitatively gauge the divergence between the optimal and Pigouvian correction in practice. Moreover, we perform comparative statics exercises, comparing optimal policies for different levels of rent-seeking, intended to capture changes in the composition of activities between the 1970s and today.

Our data source is the Current Population Survey (CPS). We take this data as generated by a (sub-optimal) tax equilibrium and use parametric assumptions and equilibrium restrictions from our model to identify the rent-seeking technology and the underlying skill distribution. Specifically, we use information on worker earnings and hours to generate a sample of hourly wages for the U.S. working population.[22] In addition, the CPS provides an industry classification that we use to assign individuals to rent-seeking versus traditional work (see Acemoglu and Autor, 2011, and Ales et al., 2014, for recent related exercises). For the sake of illustration and in the spirit of Lockwood et al. (2014), we associate industries related to finance and law services with rent-seeking and all other industries with traditional work, and consider the special case of our model where individuals always specialize in one activity by taking $m(e_\theta, e_\varphi) = e_\theta + e_\varphi$.

We assume that the observed wage distribution and sectoral choices are generated from a two-sector Roy model with individuals whose skills (θ, φ) are drawn from a bivariate lognormal distribution so that, for given E, potential wages $(\theta, \varphi\mu(E)/E)$ are also bivariate lognormal. As shown by Heckman and Honoré (1990) and French and Taber (2010), the 5 parameters of this bivariate wage distribution (2 means, 2 variances and the correlation) are identified and can be estimated by method of moments, matching the conditional means and variances in each sector as well as the share of rent-seekers in the sample. Appendix B provides further details and illustrates the quality of fit between the empirical and fitted sectoral wage distributions based on the 2013 CPS.

We assume quasilinear preferences $u(c, e) = c - h(e)$ with isoelastic disutility $h(e) = e^\gamma/\gamma$ and labor supply elasticity $1/(\gamma - 1) = 0.5$. As for the rent-seeking technology, we consider for the purpose of illustration the extreme case where $\mu(E) = \overline{\mu}$ and estimate $\overline{\mu}$ from the total income generated by rent-seekers in the calibrated economy. Of course, we do not view all of finance and law as pure rent-seeking, but take this as an illustrative extreme case for taxes, as in this case $t_{Pigou} = 100\%$. We know from Proposition 2 that $\xi < t_{Pigou}$ since $\Delta\beta = 1 > 0$ (and the other conditions will turn out to be satisfied), and will now be interested in *how much lower* the optimal correction ξ will turn out due to the relative return effects emphasized here.

In particular, we draw a large sample of potential wages (w_θ, w_φ) from the estimated bivariate distribution. From this we infer sectoral choices and wages $w = \max\{w_\theta, w_\varphi\}$. Taking the data to be generated by a $\tau = 20\%$ average tax rate (see e.g. Saez, 2001), we can back out individual efforts from $e^{\gamma-1} = (1-\tau)w$ and sum up rent-seeking incomes to obtain $\overline{\mu}$. Finally, we can, w.l.o.g., normalize $E = \overline{\mu}$. This is because scaling all φ-skills

[22] For further details on the data and sample selection, see Appendix B.

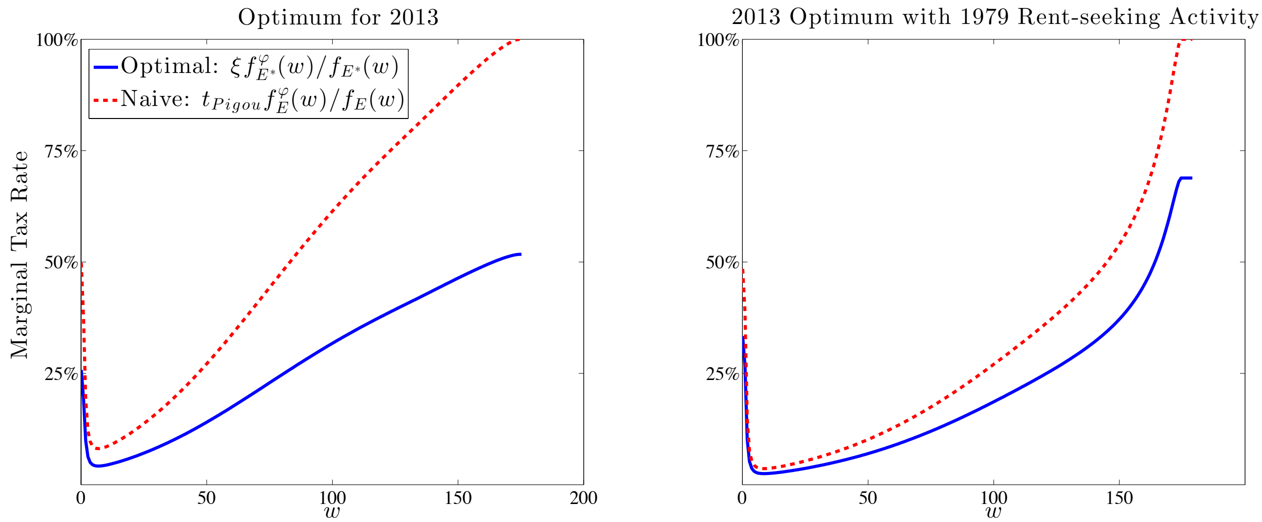

Figure 1: Optimal and weighted Pigouvian income taxes for 2013, different rent levels

by some $k > 0$ scales E by k and hence $\overline{\mu}/E$ by $1/k$, which leaves rent-seeking wages $w_\varphi = \varphi\overline{\mu}/E$, efforts and incomes unchanged. In other words, the rent-seeking skills φ are only identified up to such re-scalings for given observables.

Using this procedure, we find an overall 11% share of rent-seekers, and total rents $\overline{\mu}$ equivalent to 18% of total income. The left panel in Figure 1 shows the optimal marginal income tax for this economy (the solid blue line) under utilitarian Pareto weights $\Psi = F$, which together with quasilinear preferences imply the absence of redistributive motives. This captures the benchmark where the income tax purely serves corrective purposes and, by Proposition 1, is given simply by $T'(y(w)) = \xi f^{\varphi}_{E^*}(w)/f_{E^*}(w)$, where E^* is the optimal aggregate rent-seeking effort. We contrast this with the simple weighted average of the Pigouvian correction found to be optimal by Lockwood et al. (2014) (the dashed red line), which here is just the share of rent-seekers $f^{\varphi}_{E}(w)/f_{E}(w)$ at any given wage, since $t_{Pigou} = 1$ (and $E = \overline{\mu}$ is taken from the calibrated economy).

As can be seen, the share of rent-seekers is increasing in w for most of the wage distribution and converges to 1 for very high wages given the calibrated skill distribution.[23] The optimal policy, however, involves considerably lower marginal tax rates, since $\xi = .53$. As a result, the top marginal tax rate, for instance, is only 53% even though everyone at the top engages purely in rent-seeking. In other words, the activity shift effects $S, C > 0$ emphasized here reduce the optimal correction to roughly one-half of what a

[23] See Appendix B for a comparison to the empirical distributions. The bivariate lognormal distribution matches the empirical patterns well for intermediate wages where most of the mass is located, but overestimates the share of rent-seekers for very high wages.

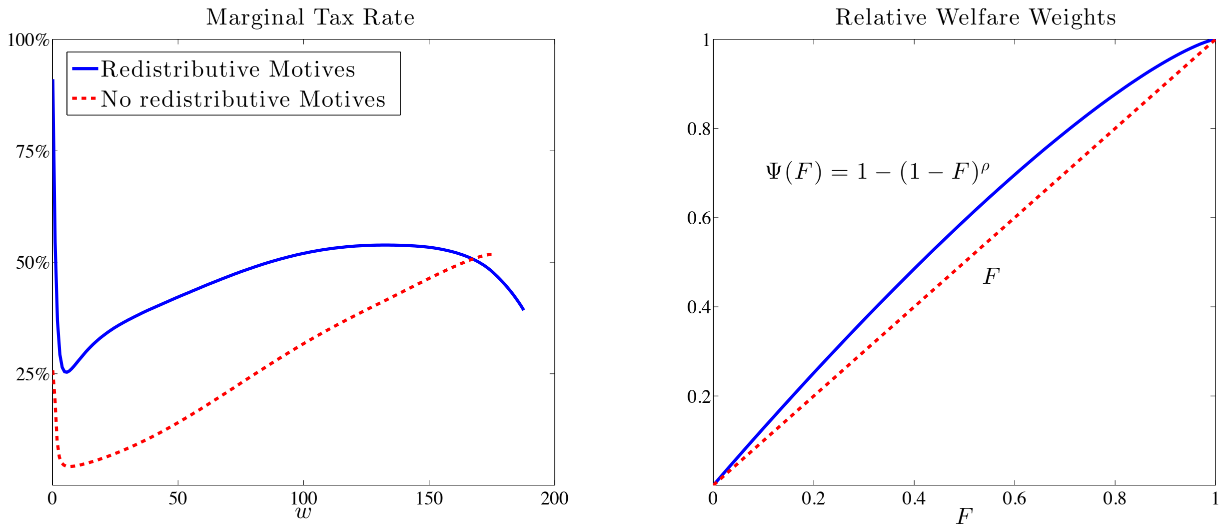

Figure 2: Optimal income taxes with and without redistributive motives

simple Pigouvian intuition would suggest, indicating that these effects can be of first-order importance, and of similar magnitude as the Pigouvian correction itself.[24]

The right panel in Figure 1 performs a counterfactual comparative statics exercise, asking: what would be the optimal policy today if the share of rent-seeking in the economy was at the level of the 1970's rather than today's? Specifically, we retain the skill distribution as identified from the 2013 data, but use CPS data from 1979 to reduce $\overline{\mu}$ so that the rent-seeking share of income in the calibrated economy is reduced to its 1979 level, roughly 6%.[25] Clearly, the reduced rents lead to a lower share of rent-seekers in the economy and hence a lower simple weighted Pigouvian correction, as indicated by the red dashed line. However, interestingly, the (solid blue) optimal marginal income tax does not fall in a proportional way: in fact, ξ now increases to .7, which shows that the relative return effects S and C also decrease when $\overline{\mu}$ falls. As a result, the divergence between the optimal and Pigouvian correction becomes smaller, and, perhaps somewhat unexpectedly, the top marginal tax rate is now actually *higher* than in the 2013 optimum.[26]

Finally, it is easy to use our general framework to include redistributive motives for

[24]Note that $I = R = 0$ under the specifications chosen here. It is easy to check numerically that effort $e(w)$ is increasing and $C > 0$. This also implies that $y(w)$ is increasing, so that bunching is not part of the optimum.

[25]All 1979 dollar variables are inflated to 2013 levels using the CPI.

[26]In Appendix B, we also show the optimal policy for 1979 itself, based on both the level of rent-seeking and the skill distribution identified from the 1979 data. These two exercises together can be interpreted as *decomposing* the change in optimal (corrective) optimal income tax rates between the 1970's and today into the component resulting from a change in rent-seeking and the component due to the change in the skill distribution. See Ales et al. (2014) for similar decomposition exercises in the face of technological change.

taxation. Figure 2 shows the optimal policy for relative Pareto weights $\Psi(F) = 1 - (1 - F)^\rho$. The case with $\rho = 1$ captures the absence of redistributive motives whereas $\rho \to \infty$ converges to a Rawlsian criterion. We pick an intermediate $\rho = 1.3$ as depicted in the right panel. The left panel shows the optimal tax policy (the blue line), contrasted with the optimum for $\rho = 1$ from Figure 1 (the red dashed line). These redistributive motives imply $I > 0$ and hence, by equation (20), generate yet another force that drives a wedge between the optimal and Pigouvian correction. Consistent with this intuition, we now find $\xi = .4$, implying an even lower top marginal tax rate than before. This indicates that accounting for redistribution can have further quantitatively large effects compared to purely efficiency-based approaches, such as Diamond (1973) in the different context of Ramsey commodity taxation, or approaches that rule out both general equilibrium effects and redistribution, such as Lockwood et al. (2014).

6 Conclusion

Our results are driven by the fact that income taxes, even when they can be conditioned on particular occupations or industries, are an imperfect instrument for correcting rent-seeking externalities. Directly taxing the externality-causing rent-seeking activity, were it possible, would reduce both its absolute desirability and its desirability relative to other activities. By contrast, an income tax for an economy or sector in which different activities take place directly affects only the absolute desirability of rent-seeking. The magnitude of the optimal correction via the income tax depends, however, on the direction of the indirect (general equilibrium) effects of taxes on the *relative* desirability of rent-seeking. When within-activity externalities dominate, these indirect effects are perverse: higher taxes on portions of the income distribution with high levels of rent-seeking raise the relative returns to rent-seeking and encourage a shift towards these activities. Consequently, the optimal externality correction lies strictly below the Pigouvian correction.

In a version of our model calibrated to U.S. data, we demonstrate how our analytical results can be operationalized for optimal tax design. The simulations imply a quantitatively important divergence between the optimal and simple Pigouvian correction. More generally, our results emphasize that the form of rent-seeking is crucial for optimal tax design: Are the returns to rent-seeking higher than its social marginal product because it depresses other rent-seekers' returns or the returns to productive activities? We view providing further quantitative evidence on these questions as an important direction for future research.

References

ACEMOGLU, D., AND D. AUTOR (2011): "Skills, Tasks and Technologies: Implications for Employment and Earnings," in *Handbook of Labor Economics*, ed. by D. Card, and O. Ashenfelter. Elsevier.

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

ALES, L., M. KURNAZ, AND C. SLEET (2014): "Tasks, Talents, and Taxes," *Mimeo, Tepper School of Business, Carnegie Mellon University*.

ARROW, K. (1962): "Economic Welfare and the Allocation of Resources for Invention," in *The Rate and Direction of Inventive Activity*, ed. by K. Arrow, pp. 609–625. Princeton University Press, Princeton NJ.

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

AXELSON, U., AND P. BOND (2012): "Wall Street Occupations: An Equilibrium Theory of Overpaid Jobs," *Mimeo, LSE*.

BAKER, G. (1992): "Incentive Contracts and Performance Measurement," *Journal of Political Economy*, 100, 98–114.

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

BEAUDRY, P., C. BLACKORBY, AND D. SZALAY (2009): "Taxes and Employment Subsidies in Optimal Redistribution Programs," *American Economic Review*, 99, 216–242.

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

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*.

CAHUC, P., AND E. CHALLE (2012): "Produce or Speculate? Asset Bubbles, Occupational Choice and Efficiency," *International Economic Review*, 53, 1105–1131.

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.

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

FRENCH, E., AND C. TABER (2010): "Identification of Models of the Labor Market," *Federal Reserve Bank of Chicago Working Paper 2010-08*.

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., P. MAZIERO, AND G. MENZIO (2013): "Taxation and Redistribution of Residual Income Inequality," *Journal of Political Economy*, 121, 1160–1204.

HEATHCOTE, J., F. PERRI, AND G. VIOLANTE (2010): "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967-2006," *Review of Economic Dynamics*, 13, 15–51.

HECKMAN, J., AND B. HONORE (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58, 1121–1149.

HIRSHLEIFER, J. (1971): "The private and social value of information and the reward to inventive activity," *American Economic Review*, 61, 561–574.

HOLMSTRÖM, B., AND P. MILGROM (1991): "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, 7, 24–52.

HUNGERBUHLER, M., E. LEHMANN, A. PARMENTIER, AND B. LINDEN (2008): "Optimal redistributive taxation in a search equilibrium model," *Review of Economic Studies*, 73, 743–767.

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.

LAZEAR, E., AND S. ROSEN (1979): "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841–864.

LOCKWOOD, B., C. NATHANSON, AND G. WEYL (2014): "Taxation and the Allocation of Talent," *Mimeo, Harvard University and University of Chicago*.

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

MURPHY, K., A. SHLEIFER, AND R. VISHNY (1991): "The Allocation of Talent: Implications for Growth," *Quarterly Journal of Economics*, 106, 503–530.

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

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.

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*.

——— (2013): "Redistributive Taxation in the Roy Model," *Quarterly Journal of Economics*, 128, 623–668.

——— (2014): "A Theory of Income Taxation under Multidimensional Skill Heterogeneity," *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.

SAEZ, E., J. SLEMROD, AND S. GIERTZ (2012): "The Elasticity of Taxable Income with Respect to the Marginal Tax Rate," *Journal of Economic Literature*, 50.

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

SCHEUER, F. (2013): "Adverse Selection in Credit Markets and Regressive Profit Taxation," *Journal of Economic Theory*, 148, 1333–1360.

——— (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.

TULLOCK, G. (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 4

A.1 Proof of Proposition 1

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

$$\begin{aligned}\mathcal{L} &= \textstyle\int_{\underline{w}_E}^{\overline{w}_E} V(w)\psi_E(w)dw - \int_{\underline{w}_E}^{\overline{w}_E} V(w)\hat{\eta}'(w)\lambda dw + \int_{\underline{w}_E}^{\overline{w}_E} u_e(c(V(w),e(w)),e(w))\frac{e(w)}{w}\hat{\eta}(w)\lambda dw \\ &+\xi\lambda\mu(E) - \xi\lambda \textstyle\int_{\underline{w}_E}^{\overline{w}_E} we(w)f_E^{\varphi}(w)dw + \lambda \int_{\underline{w}_E}^{\overline{w}_E} we(w)f_E(w)dw - \lambda \int_{\underline{w}_E}^{\overline{w}_E} c(V(w),e(w))f_E(w)dw. \quad (24)\end{aligned}$$

By Theorem 3 in Clarke (1976), this Lagrangian approach is generically valid, i.e. constraint qualification is satisfied and any solution must be a stationary point of Lagrangian for some (bounded) multipliers. Using $\partial c/\partial V = 1/u_c$ and compressing notation, the first order condition for $V(w)$ is

$$\hat{\eta}'(w)\lambda = \psi_E(w) - \lambda f_E(w)\frac{1}{u_c(w)} + \hat{\eta}(w)\lambda\frac{u_{ec}(w)}{u_c(w)}\frac{e(w)}{w}. \quad (25)$$

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

$$\eta'(w) = \psi_E(w)\frac{u_c(w)}{\lambda} - f_E(w) + \eta(w)\frac{u_{cc}(w)c'(w) + u_{ce}(w)e'(w) + u_{ce}(w)e(w)/w}{u_c(w)}. \quad (26)$$

Using the first order condition corresponding to the incentive constraint (10),

$$u_c(w)c'(w) + u_e(w)e'(w) + u_e(w)\frac{e(w)}{w} = 0, \quad (27)$$

the fraction in (26) can be written as $-(\partial MRS(w)/\partial c)y'(w)/w$, where $M(c,e) \equiv -u_e(c,e)/u_c(c,e)$ is the marginal rate of substitution between effort and consumption and $MRS(w) \equiv M(c(w),e(w))$, so (with a slight abuse of notation) $\partial MRS(w)/\partial c$ stands short for $\partial M(c(w),e(w))/\partial c$. Substituting in (26) and rearranging yields

$$-\frac{\partial MRS(w)}{\partial c}e(w)\frac{y'(w)}{y(w)}\eta(w) = f_E(w) - \psi_E(w)\frac{u_c(w)}{\lambda} + \eta'(w). \quad (28)$$

Integrating this ODE gives

$$\begin{aligned}\eta(w) &= \int_w^{\overline{w}_E}\left(f_E(x) - \psi_E(x)\frac{u_c(x)}{\lambda}\right)\exp\left(\int_w^x \frac{\partial MRS(s)}{\partial c}e(s)\frac{y'(s)}{y(s)}ds\right)dx \\ &= \int_w^{\overline{w}_E}\left(1 - \frac{\psi_E(x)}{f_E(x)}\frac{u_c(x)}{\lambda}\right)\exp\left(\int_w^x\left(1 - \frac{\varepsilon^u(s)}{\varepsilon^c(s)}\right)\frac{dy(s)}{y(s)}\right)f_E(x)dx, \quad (29)\end{aligned}$$

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

$$wf_E(w)\left(1-\frac{MRS(w)}{w}\right)-\xi wf_E^{\varphi}(w)=\eta(w)\left(\frac{\partial MRS(w)}{\partial e}\frac{e}{w}+\frac{MRS(w)}{w}\right), \tag{30}$$

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

$$1-\xi\frac{f_E^{\varphi}(w)}{f_E(w)}=(1-T'(y(w)))\left[1+\frac{\eta(w)}{wf_E(w)}\left(1+\frac{\partial MRS(w)}{\partial e}\frac{e}{MRS(w)}\right)\right]. \tag{31}$$

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 (29) and (31).

A.2 Proof of Lemma 3

Because we allow m to be only weakly quasiconvex, the marginal rate of substitution $\chi(e_\theta/e_\varphi)\equiv m_\theta(e_\theta/e_\varphi)/m_\varphi(e_\theta/e_\varphi)$ can have constant regions. Since χ is non-decreasing, there are at most a countable number of such regions. Use $\{s_i\}_{i=1}^n$, with s_i increasing in i, to denote the values of χ at these flat regions.[27] An individual with $\theta/\varphi=\frac{\mu(E)}{E\Gamma(E)}s_i\equiv\alpha_E^i$ will be indifferent between a range of activity-specific income shares $q\in\left[\underline{q}^i,\overline{q}^i\right]$, where, by equation (5), the bounds are, respectively, the minimum and the maximum q for which $\chi\left(\frac{q}{1-q}\frac{1}{s_i}\right)=s_i$ and thus independent of E. The upper-hemicontinuous correspondence $q_E(\alpha)$ thus jumps from $\underline{q}^i$ to $\overline{q}^i$ as α crosses α_E^i from below. For any given wage w and E, taking $\alpha_E^0=\underline{\theta}/\overline{\varphi}$ and $\alpha_E^{n+1}=\overline{\theta}/\underline{\varphi}$, we can thus write[28]

$$F_E(w)=\int_{\alpha_E^0}^{\alpha_E^{n+1}}\int_{\underline{w}_E(\alpha)}^{w}g_E(w',\alpha)dw'd\alpha=\sum_{i=0}^{n}\int_{\alpha_E^i}^{\alpha_E^{i+1}}\int_{\underline{w}_E(\alpha)}^{w}g_E(w',\alpha)dw'd\alpha,$$

where $\underline{w}_E(\alpha)\equiv w_E\left(\max\{\underline{\theta},\alpha\underline{\varphi}\},\max\{\underline{\theta}/\alpha,\underline{\varphi}\}\right)$ and $\overline{w}_E(\alpha)\equiv w_E\left(\min\{\overline{\theta},\alpha\overline{\varphi}\},\min\{\overline{\theta}/\alpha,\overline{\varphi}\}\right)$. Similarly, we can write

$$F_E^{\theta}(w)=\sum_{i=0}^{n}\int_{\alpha_E^i}^{\alpha_E^{i+1}}\int_{\underline{w}_E(\alpha)}^{w}q_E(\alpha)g_E(w',\alpha)dw'd\alpha.$$

The latter is a useful formulation because, on each of the intervals in the sum, the function $q_E(\alpha)$ is continuously differentiable. The discontinuities occur at the boundaries of the intervals. It allows us to prove the following technical lemma, which will be useful below.

Lemma 4.

$$\frac{dF_E(w)}{dE}=\frac{\beta_\varphi(E)}{E}wf_E^{\varphi}(w)+\frac{\beta_\theta(E)}{E}wf_E^{\theta}(w)\ \ and\ \ \frac{dF_E^{\theta}(w)}{dE}=\frac{\beta_\theta(E)}{E}wf_E^{\theta}(w)-K_E(w)-L_E(w),\ where \tag{32}$$

[27] We deal with finite n to keep notation clean; the countably infinite case is an easy but notationally cumbersome extension.

[28] We mildly abuse notation here in assuming that $\alpha_E^0\leq\alpha_E^1$ and $\alpha_E^{n+1}\geq\alpha_E^n$ since a simple redefinition of n and the relevant α_E^i values would yield the same formulae.

$$
\begin{aligned}
K_E(w) &\equiv -\sum_{i=0}^{n}\int_{\alpha_E^i}^{\alpha_E^{i+1}}\int_{\underline{w}_E(\alpha)}^{w}\frac{dq_E(\alpha)}{dE}g_E(w',\alpha)dw'd\alpha+\sum_{i=1}^{n}\int_{\underline{w}_E(\alpha_E^i)}^{w}\frac{d\alpha_E^i}{dE}\left(\overline{q}^i-\underline{q}^i\right)g_E\left(w',\alpha_E^i\right)dw' \\
&= -\frac{\Delta\beta(E)}{E}\int_{\alpha_E^0}^{\alpha_E^{n+1}}\int_{\underline{w}_E(\alpha)}^{w}Q'(x_E(\alpha))x_E(\alpha)g_E(w',\alpha)dw'd\alpha \qquad (33)
\end{aligned}
$$

$$
\textit{and}\quad L_E(w)\equiv-\frac{\Delta\beta(E)}{E}\int_{\alpha_E^0}^{\alpha_E^{n+1}}q_E(\alpha)(1-q_E(\alpha))wg_E(w,\alpha)d\alpha. \qquad (34)
$$

Moreover, $dF_E^{\varphi}(w)/dE=dF_E(w)/dE-dF_E^{\theta}(w)/dE$ *and analogous expressions hold for* $d\Psi_E(w)/dE$, $d\Psi_E^{\theta}(w)/dE$ *and* $d\Psi_E^{\varphi}(w)/dE$.

Proof. It is useful to define the function $\tilde{w}(w,\alpha,E;E_0)$ as the wage, at E, of the type (θ,φ) that would have had wage w (and α) at $E=E_0$. Then, by construction, the set $\{(\theta,\varphi)|w_E(\theta,\varphi)\le\tilde{w}(w,\theta/\varphi,E;E_0)\}$ is independent of E: it is simply the set of types that would have had wage less than w at E_0. Hence, for all E,

$$
F_{E_0}(w)=\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}}\int_{\underline{w}_E(\alpha)}^{\tilde{w}(w,\alpha,E;E_0)}g_E(w',\alpha)dw'd\alpha \qquad (35)
$$

i.e., the measure of types with wage less than w when $E=E_0$. Taking the derivative of the RHS with respect to E (which is zero by construction) and evaluating at E_0 yields

$$
\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}}\int_{\underline{w}_E(\alpha)}^{w}\frac{dg_{E_0}(w',\alpha)}{dE}dw'd\alpha-\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}}\frac{d\underline{w}_{E_0}(\alpha)}{dE}g_{E_0}(\underline{w}_{E_0},\alpha)d\alpha=-\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}}\frac{d\tilde{w}(w,\alpha,E_0;E_0)}{dE}g_{E_0}(w,\alpha)d\alpha.
$$

The LHS is easily recognized as $dF_E(w)/dE$, evaluated at E_0. Using (19) to obtain $d\tilde{w}(w,\alpha,E_0;E_0)/dE$,

$$
\frac{dF_E(w)}{dE}=\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}}\left(q_E(\alpha)\frac{\beta_\theta(E)}{E}w+(1-q_E(\alpha))\frac{\beta_\varphi(E)}{E}w\right)g_E(w,\alpha)d\alpha=\frac{\beta_\theta(E)}{E}wf_E^{\theta}(w)+\frac{\beta_\varphi(E)}{E}wf_E^{\varphi}(w),
$$

which proves the first result in (32). Moreover, for all E,

$$
F_{E_0}^{\theta}(w)=\sum_{i=0}^{n}\int_{\alpha_{E_0}^i}^{\alpha_{E_0}^{i+1}}\int_{\underline{w}_E(\alpha)}^{\tilde{w}(w,\alpha,E;E_0)}q_{E_0}(\alpha)g_E(w',\alpha)dw'd\alpha, \qquad (36)
$$

since the set of types being integrated over is independent of E, and so is $q_{E_0}(\alpha)$. We explicitly compute the derivative with respect to E of the RHS of (36), which is zero since the object is, by construction, independent of E. After some re-arranging and evaluating at E_0, we get

$$
\begin{aligned}
&\sum_{i=0}^{n}\int_{\alpha_{E_0}^i}^{\alpha_{E_0}^{i+1}}\int_{\underline{w}_{E_0}(\alpha)}^{w}q_{E_0}(\alpha)\frac{dg_{E_0}(w',\alpha)}{dE}dw'd\alpha-\sum_{i=0}^{n}\int_{\alpha_{E_0}^i}^{\alpha_{E_0}^{i+1}}\frac{d\underline{w}_{E_0}(\alpha)}{dE}q_{E_0}(\alpha)g_{E_0}\left(\underline{w}_{E_0}(\alpha),\alpha\right)d\alpha \\
&=-\sum_{i=0}^{n}\int_{\alpha_{E_0}^i}^{\alpha_{E_0}^{i+1}}\frac{d\tilde{w}(w,\alpha,E_0;E_0)}{dE}q_{E_0}(\alpha)g_{E_0}(w,\alpha)d\alpha. \qquad (37)
\end{aligned}
$$

Adding $\sum_{i=0}^{n}\int_{\alpha_{E_0}^i}^{\alpha_{E_0}^{i+1}}\int_{\underline{w}_{E_0}(\alpha)}^{w}\frac{dq_{E_0}(\alpha)}{dE}g_{E_0}(w',\alpha)dw'd\alpha-\sum_{i=1}^{n}\int_{\underline{w}_{E_0}(\alpha_{E_0}^i)}^{w}\frac{d\alpha_{E_0}^i}{dE}\left(\overline{q}^i-\underline{q}^i\right)g_{E_0}\left(w',\alpha_{E_0}^i\right)dw'$ to both sides, we can recognize the left hand side as $dF^{\theta}_E(w)/dE$ evaluated at E_0. Again using (19), the RHS of (37) is

$$\int_{\alpha_E^0}^{\alpha_E^{n+1}} \left(q_{E_0}(\alpha)\frac{\beta_\theta(E_0)}{E_0}w + (1-q_{E_0}(\alpha))\frac{\beta_\varphi(E_0)}{E_0}w \right) q_{E_0}(\alpha) g_{E_0}(w,\alpha)d\alpha$$
$$= \frac{\beta_\theta(E_0)}{E_0} w f^{\theta}_{E_0}(w) + \frac{\Delta\beta(E_0)}{E_0}\int_{\alpha_E^0}^{\alpha_E^{n+1}} q_{E_0}(\alpha)(1-q_{E_0}(\alpha)) w g_{E_0}(w,\alpha)d\alpha.$$

Using the definitions in (33) and (34), we conclude that $dF^{\theta}_E(w)/dE = \frac{\beta_\theta(E)}{E} w f^{\theta}_E(w) - K_E(w) - L_E(w)$.

Finally, observe that $\tilde{q}_E(\alpha) \equiv q_E(\alpha) - \sum_{i=1}^{n}(\overline{q}^i - \underline{q}^i)H(\alpha - \alpha_E^i)$, where H is the Heaviside step function (using the half-maximum convention), is continuous. It has a well-defined derivative with respect to E equal to $dq_E(\alpha)/dE$ away from the α_E^i-jumps, and well-defined (bounded) left- and right-derivatives at the jumps. Therefore,

$$\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}} \int_{\underline{w}_E(\alpha)}^{w} \frac{d\tilde{q}_E(\alpha)}{dE} g_E(w',\alpha)dw'd\alpha = \sum_{i=0}^{n}\int_{\alpha=\alpha_E^i}^{\alpha_E^{i+1}} \int_{\underline{w}_E(\alpha)}^{w} \frac{dq_E(\alpha)}{dE} g_E(w',\alpha)dw'd\alpha. \tag{38}$$

Differentiating H as a distribution yields $\frac{dH(\alpha-\alpha_E^i)}{dE} = -\delta(\alpha-\alpha_E^i)\frac{d\alpha_E^i}{dE}$, where δ is the Dirac δ-function. Hence,

$$\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}} \int_{\underline{w}_E(\alpha)}^{w} (\overline{q}^i - \underline{q}^i)\frac{dH(\alpha-\alpha_E^i)}{dE} g_E(w',\alpha)dw'd\alpha = -\int_{\underline{w}_E(\alpha_E^i)}^{w} \frac{d\alpha_E^i}{dE}(\overline{q}^i - \underline{q}^i) g_E(w',\alpha_E^i)dw'. \tag{39}$$

Combining (38) and (39) (and differentiating $q_E(\alpha)$ as a distribution), we have

$$\int_{\alpha=\alpha_E^0}^{\alpha_E^{n+1}} \int_{\underline{w}_E(\alpha)}^{w} \frac{dq_E(\alpha)}{dE} g_E(w',\alpha)dw'd\alpha = \sum_{i=0}^{n}\int_{\alpha=\alpha_E^i}^{\alpha_E^{i+1}} \int_{\underline{w}_E(\alpha)}^{w} \frac{dq_E(\alpha)}{dE} g_E(w',\alpha)dw'd\alpha - \sum_{i=0}^{n}\int_{\underline{w}_E(\alpha_E^i)}^{w} \frac{d\alpha_E^i}{dE}(\overline{q}^i - \underline{q}^i) g_E(w',\alpha_E^i)dw'.$$

Using $Q(x_E(\alpha)) \equiv q_E(\alpha)$ and $dx_E(\alpha)/dE = x_E(\alpha)\Delta\beta(E)/E$ yields the second equality in (33). □

Lemma 4 will be used in the now following proof of Lemma 3. Using (24) and a standard envelope theorem,

$$W'(E) = \int_{\underline{w}_E}^{\overline{w}_E} V(w)\frac{d\psi_E(w)}{dE}dw - \lambda\int_{\underline{w}_E}^{\overline{w}_E} c(V(w),e(w))\frac{df_E(w)}{dE}dw$$
$$+ \lambda(1-\xi)\int_{\underline{w}_E}^{\overline{w}_E} we(w)\frac{df^{\varphi}_E(w)}{dE}dw + \lambda\int_{\underline{w}_E}^{\overline{w}_E} we(w)\frac{df^{\theta}_E(w)}{dE}dw + \xi\lambda\mu'(E) + B_1 \tag{40}$$

with

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

Integrating by parts the four integrals yields

$$W'(E) = B_1 + B_2 - \int_{\underline{w}_E}^{\overline{w}_E} V'(w)\frac{d\Psi_E(w)}{dE}dw + \lambda\int_{\underline{w}_E}^{\overline{w}_E}\left(\frac{V'(w)}{u_c(w)} + MRS(w)e'(w)\right)\frac{dF_E(w)}{dE}dw$$
$$-\lambda(1-\xi)\int_{\underline{w}_E}^{\overline{w}_E}(we'(w)+e(w))\frac{dF^{\varphi}_E(w)}{dE}dw - \lambda\int_{\underline{w}_E}^{\overline{w}_E}(we'(w)+e(w))\frac{dF^{\theta}_E(w)}{dE}dw + \xi\lambda\mu'(E) \tag{41}$$

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 (28) and (30) with respect to $V(w)$ and $e(w)$ from the inner problem, the terms

$$A_1(E) \equiv \lambda \int_{\underline{w}_E}^{\overline{w}_E} \frac{e'(w)}{f_E(w)} \left[wf_E(w)\left(1 - \frac{MRS(w)}{w}\right) - \xi w f_E^{\varphi}(w) - \eta(w)\left(\frac{\partial MRS(w)}{\partial e}\frac{e(w)}{w} + \frac{MRS(w)}{w}\right)\right] \frac{dF_E(w)}{dE} dw$$

and $$A_2(E) \equiv \lambda \int_{\underline{w}_E}^{\overline{w}_E} \frac{V'(w)}{u_c(w) f_E(w)} \left[\psi_E(w)\frac{u_c(w)}{\lambda} - f_E(w) - \eta'(w) - \eta(w)\frac{\partial MRS(w)}{\partial c} e(w)\frac{y'(w)}{y(w)}\right] \frac{dF_E(w)}{dE} dw$$

are both equal to zero. Adding $A_1(E)$ and $A_2(E)$ to (41), using (7) and re-arranging yields

$$\begin{aligned} W'(E) = {} & B_1 + B_2 + \xi\lambda\mu'(E) + \int_{\underline{w}_E}^{\overline{w}_E} V'(w)\left(\frac{\psi_E(w)}{f_E(w)}\frac{dF_E(w)}{dE} - \frac{d\Psi_E(w)}{dE}\right) dw \\ & - \lambda \int_{\underline{w}_E}^{\overline{w}_E} e(w)\frac{dF_E(w)}{dE} dw + \xi\lambda \int_{\underline{w}_E}^{\overline{w}_E} \left((e(w) + we'(w))\frac{dF_E^{\varphi}(w)}{dE} - we'(w)\frac{f_E^{\varphi}(w)}{f_E(w)}\frac{dF_E(w)}{dE}\right) dw \\ & - \lambda \int_{\underline{w}_E}^{\overline{w}_E} \left(\frac{\eta(w)}{w}\frac{d\left[MRS(w)e(w)\right]}{dw} + \eta'(w)\frac{V'(w)}{u_c(w)}\right)\frac{1}{f_E(w)}\frac{dF_E(w)}{dE} dw. \end{aligned} \tag{42}$$

Using Lemma 4, the first integral in (42) is

$$\frac{\Delta\beta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} V'(w) w \frac{f_E^{\theta}(w) f_E^{\varphi}(w)}{f_E(w)}\left(\frac{\psi_E^{\theta}(w)}{f_E^{\theta}(w)} - \frac{\psi_E^{\varphi}(w)}{f_E^{\varphi}(w)}\right) dw = \frac{\Delta\beta(E)}{E} R \tag{43}$$

Again using Lemma 4, the terms with $e(w)$ in the second line of (42) can be written as

$$\begin{aligned} & -\lambda(1-\xi)\frac{\beta_{\varphi}(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} we(w) f_E^{\varphi}(w) dw - \lambda\frac{\beta_{\theta}(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} we(w) f_E^{\theta}(w) dw + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} e(w)(K_E(w) + L_E(w)) dw \\ & = -\lambda\left(\beta_{\varphi}(E)\frac{\mu(E)}{E} + \beta_{\theta}(E)\frac{\Gamma(E)}{E}E_{\theta}\right) + \xi\lambda\beta_{\varphi}(E)\frac{\mu(E)}{E} + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} e(w)(K_E(w) + L_E(w)) dw \\ & = -\lambda\frac{\mu(E)}{E}t_{Pigou} + \xi\lambda\beta_{\varphi}(E)\frac{\mu(E)}{E} + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} e(w)(K_E(w) + L_E(w)) dw. \end{aligned} \tag{44}$$

The terms with $we'(w)$ in (42) can be written as

$$\begin{aligned} & \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} we'(w)\left[\frac{\beta_{\varphi}(E)}{E} w f_E^{\varphi}(w) - \frac{f_E^{\varphi}(w)}{f_E(w)}\frac{w}{E}\left(\beta_{\varphi}(E) f_E^{\varphi}(w) + \beta_{\theta}(E) f_E^{\theta}(w)\right) + K_E(w) + L_E(w)\right] dw \\ = {} & \xi\lambda\frac{\Delta\beta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} w^2 e'(w)\left[\frac{f_E^{\theta}(w) f_E^{\varphi}(w)}{f_E(w)} - \int_0^{\infty} q_E(\alpha)(1 - q_E(\alpha)) g_E(w,\alpha) d\alpha\right] dw + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} we'(w) K_E(w) dw \\ = {} & \xi\lambda\frac{\Delta\beta(E)}{E}\int_{\underline{w}_E}^{\overline{w}_E} w^2 e'(w) \mathrm{Var}_E(q|w) f_E(w) dw + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} we'(w) K_E(w) dw \\ = {} & \xi\lambda\frac{\Delta\beta(E)}{E} C + \xi\lambda\int_{\underline{w}_E}^{\overline{w}_E} we'(w) K_E(w) dw, \end{aligned} \tag{45}$$

where the first equality uses (34). Combining the terms with $K_E(w)$ from (44) and (45) gives $\xi\lambda\int_{\underline{w}_E}^{\overline{w}_E}(we(w))' K_E(w) dw$, which can be integrated by parts to yield

$$B_3 - \xi\lambda \int_{\underline{w}_E}^{\overline{w}_E} we(w)K'_E(w)dw = B_3 + \xi\lambda \frac{\Delta\beta(E)}{E} \int_{\underline{w}_E}^{\overline{w}_E} \int_0^\infty y(w)Q'(x_E(\alpha))x_E(\alpha)dG_E(w,\alpha) \tag{46}$$

with $B_3 = \xi\lambda\overline{w}_E e(\overline{w}_E)K_E(\overline{w}_E)$ since $K_E(\underline{w}_E) = 0$. Further combining this with the $L_E(w)$-term in (44) yields

$$B_3 + \xi\lambda \frac{\Delta\beta(E)}{E} \int_{\underline{w}_E}^{\overline{w}_E} y(w) \int_0^\infty \left(Q'(x_E(\alpha))x_E(\alpha) - q_E(\alpha)(1 - q_E(\alpha))\right) dG_E(w,\alpha). \tag{47}$$

Note that $\tilde{Q}'(x) = \tilde{Q}'_{x_0}(x)\big|_{x_0=x} = Q'(x)x - Q(x)(1 - Q(x))$, so that (47) becomes

$$B_3 + \xi\lambda \frac{\Delta\beta(E)}{E} \int_{\underline{w}_E}^{\overline{w}_E} \int_0^\infty y(w)\tilde{Q}'(x_E(\alpha))x_E(\alpha)dG_E(w,\alpha) = B_3 + \xi\lambda \frac{\Delta\beta(E)}{E} S. \tag{48}$$

Moreover, $S \geq 0$ since $\tilde{Q}'(x) = x^2 r'(x)/(1 + xr(x))^2 \geq 0$. Finally, use the incentive constraint (7), rewritten as $V'(w)/u_c(w) = MRS(w)e(w)/w$, to write the last line of (42) as

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

or, recognizing the sum of the bracketed terms as $d[\eta(w)wV'(w)/u_c(w)]/dw$, integrating by parts, and using the transversality condition $\eta(\underline{w}_E) = \eta(\overline{w}_E) = 0$ and (32),

$$\begin{aligned} &\lambda \int_{\underline{w}_E}^{\overline{w}_E} \eta(w)w \frac{V'(w)}{u_c(w)} \frac{d}{dw} \left(\frac{\beta_\theta(E)}{E} \frac{f_E^\theta(w)}{f_E(w)} + \frac{\beta_\varphi(E)}{E} \frac{f_E^\varphi(w)}{f_E(w)}\right) dw \\ &\qquad = \lambda \frac{\Delta\beta(E)}{E} \int_{\underline{w}_E}^{\overline{w}_E} \eta(w)w \frac{V'(w)}{u_c(w)} \frac{d}{dw} \left(\frac{f_E^\varphi(w)}{f_E(w)}\right) dw = \frac{\Delta\beta(E)}{E} I. \end{aligned} \tag{49}$$

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

$$\frac{d\tilde{F}(\overline{w}_E, E)}{dE} = \frac{\partial\tilde{F}(\overline{w}_E, E)}{\partial E} + \frac{\partial\tilde{F}(\overline{w}_E, E)}{\partial w} \frac{d\overline{w}_E}{dE} = \frac{dF_E(\overline{w}_E)}{dE} + f_E(\overline{w}_E) \frac{d\overline{w}_E}{dE} = 0. \tag{50}$$

Together with an analogous expression at $\underline{w}_E$ and the fact that $K_E(\underline{w}_E) = L_E(\underline{w}_E) = L_E(\overline{w}_E) = 0$, this yields $B_1 + B_2 = -\xi\lambda\overline{w}_E\, e(\overline{w}_E)K_E(\overline{w}_E) = -B_3$. Using (43), (44), (45), (48) and (49) in (42) yields

$$W'(E) = -\lambda \frac{\mu(E)}{E} t_{Pigou} + \frac{\Delta\beta(E)}{E} (R + I) + \xi\lambda \left(\frac{\mu(E)}{E} + \frac{\Delta\beta(E)}{E} (C + S)\right), \tag{51}$$

where we have used $\mu'(E) + \beta_\varphi(E)\mu(E)/E = \mu(E)/E$.

A.3 Proof of Corollary 1

Using the notation of Lemma 3, when m is linear, then $n = 1$, $\overline{q}^1 = 1$, $\underline{q}^1 = 0$, $s_1 = 1$, and $\alpha^1_E = \frac{\mu(E)}{E\Gamma(E)}$. Hence, for the region where $q = 1$,

$$K_E(w) = \sum_{i=1}^{n} \int_{\underline{w}_E}^{w} \frac{d\alpha^i_E}{dE}(\overline{q}^i - \underline{q}^i) g_E(w', \alpha^i_E) dw' = \frac{d}{dE}\left(\frac{\mu(E)}{E\Gamma(E)}\right) \int_{\underline{w}_E}^{w} g_E\left(w', \frac{\mu(E)}{E\Gamma(E)}\right) dw'.$$

In this region, $w = \theta\Gamma(E)$ and $\alpha = \theta/\varphi$. Hence, $\theta = w/\Gamma(E)$, $\varphi = \frac{w}{\alpha\Gamma(E)}$ and the Jacobian for the transformation is $\frac{w}{\alpha^2\Gamma(E)^2}$. Therefore, $g_E\left(w, \frac{\mu(E)}{E\Gamma(E)}\right) = \frac{w}{(\alpha^1_E)^2\Gamma(E)^2} f\left(\frac{w}{\Gamma(E)}, \frac{w}{\alpha^1_E\Gamma(E)}\right)$, or, transforming to $\varphi = wE/\mu(E)$,

$$K_E(w) = \frac{d}{dE}\left(\frac{\mu(E)}{E\Gamma(E)}\right) \int_{\underline{w}_E}^{w} g_E\left(w', \frac{\mu(E)}{E\Gamma(E)}\right) dw' = -\frac{\Delta\beta(E)}{E} \frac{\mu(E)}{E\Gamma(E)} \int_{\underline{\varphi}}^{wE/\mu(E)} \varphi f\left(\varphi \frac{\mu(E)}{\Gamma(E)E}, \varphi\right) d\varphi.$$

Moreover, $L_E(w) = 0$ since $q \in \{0, 1\}$. Using this in (45) and (46) yields the result.

A.4 Proof of Proposition 2

The first part of the proposition is established by the following lemma:

Lemma 5. *$\xi > 0$ in any regular Pareto optimum.*

Proof. By Lemma **??**, any Pareto optimum solves $\max_{E,T(.)} \int_{\underline{w}_E}^{\overline{w}_E} u\left(y_T(w) - T(y_T(w)), \frac{y_T(w)}{w}\right) d\Psi_E(w)$ subject to (a) a set of incentive constraints $y_T(w) \in \arg\max_y u(y - T(y), y/w)$ for all w, (b) the consistency constraint $\mu(E) - \int_{\underline{w}_E}^{\overline{w}_E} y_T(w) f^{\varphi}_E(w) dw = 0$ and (c) the revenue constraint $\int_{\underline{w}_E}^{\overline{w}_E} T(y_T(w)) dF_E(w) \geq 0$. We show that any solution to the relaxed problem where the consistency constraint (c) is replaced by the inequality constraint $\mu(E) - \int_{\underline{w}_E}^{\overline{w}_E} y_T(w) f^{\varphi}_E(w) dw \geq 0$, with associated multiplier $\lambda\tilde{\xi}$, always has $\tilde{\xi} > 0$ with regular Ψ. Since this implies that the solution $(E^*, T^*(.))$ to the relaxed problem is feasible in the unrelaxed problem, $\xi = \tilde{\xi} > 0$ as well. To see this, suppose, by way of contradiction, that $\tilde{\xi} = 0$ in $(E^*, T^*(.))$. Standard arguments (e.g. Werning, 2000) imply $T^{*\prime}(.) \geq 0$ with regular Pareto weights in this case. Now consider a small decrease ΔE from E^* holding $T^*(.)$ fixed. This at least weakly increases the wage, and hence the utility, of each individual, increasing the objective. It has no effect on the set of incentive constraints (a) since $T^*(.)$ remains fixed. It has no effect on the relaxed constraint (b) since $\tilde{\xi} = 0$. It relaxes the revenue constraint (c) since $y_T(w)$ is non-decreasing and $T^{*\prime}(.) \geq 0$. This contradicts the optimality of $(E^*, T^*(.))$ in the relaxed problem, showing that $\tilde{\xi} > 0$ in the relaxed problem and hence $\xi = \tilde{\xi} > 0$. □

We next show that $\eta(w) \geq 0$ 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 (12), 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 (20) or both are positive. $\xi > 0$ implies that both are positive. Hence, $\xi \lesseqgtr t_{Pigou} \Leftrightarrow \Delta\beta(E) \gtreqless 0$.

A.5 Proof of Proposition 3

Part (i). Consider the change in $W(E,t)$ induced by a small variation from (E,t) to $(E+\delta E, t+\delta t)$ with $\delta t = -(1-t)\Delta\beta(E)\delta E/E$. This variation is leaves $(1-t)\mu(E)/(E\Gamma(E))$ unchanged and therefore affects neither the optimal q of any individual nor

$$\tilde{\theta}(\theta,\varphi;E,t) \equiv \max_{q\in[0,1]} m\left(\frac{q}{\theta}, \frac{1-q}{(1-t)\mu(E)/(E\Gamma(E))\varphi}\right)^{-1}.$$

By the envelope theorem, the welfare effects of this variation can be computed by holding fixed the allocations $e(\tilde{\theta}), c(\tilde{\theta})$ for each individual. It therefore has no direct effect on individuals' utilities, and, since it leaves unchanged the ratio of after-t wages between any two types, does not affect the incentive constraints. To compute the effects on the consistency and resource constraints (22) and (23), it is convenient to change variables within the integrals from after-t wages w to the re-scaled wages $\tilde{\theta} \equiv w/\Gamma(E)$:[29]

$$\int_{\underline{\tilde{\theta}}_{E,t}}^{\overline{\tilde{\theta}}_{E,t}} \tilde{\theta}\Gamma(E)e(\tilde{\theta})\tilde{f}^{\varphi}_{E,t}(\tilde{\theta})d\tilde{\theta} + t\mu(E) = \mu(E) \tag{52}$$

$$\int_{\underline{\tilde{\theta}}_{E,t}}^{\overline{\tilde{\theta}}_{E,t}} \left(\Gamma(E)\tilde{\theta}e(\tilde{\theta}) - c(V(\tilde{\theta}), e(\tilde{\theta}))\right)\tilde{f}_{E,t}(\tilde{\theta})d\tilde{\theta} + (t-s)\mu(E) + s\left(\frac{\Gamma(E)E}{(1-t)\mu(E)}\right)\frac{\mu(E)}{E}\int_{\underline{\tilde{\theta}}_{E,t}}^{\overline{\tilde{\theta}}_{E,t}} \tilde{\theta}e(\tilde{\theta})\tilde{f}^{\varphi}_{E,t}(\tilde{\theta})d\tilde{\theta} = 0, \tag{53}$$

where $\tilde{f}^k_{E,t}(\tilde{\theta}) = \tilde{f}^k_{E,t}(w/\Gamma(E)) \equiv \Gamma(E)f^k_{E,t}(w)$, $k = \theta, \varphi$, are the wage densities in terms of $\tilde{\theta}$. The variation leaves everything in these constraints unchanged except the terms $\Gamma(E)$, $\mu(E)$, and t. Hence,

$$\begin{aligned}\delta W &= \lambda\left[\xi\left((1-t)\mu'(E) - \frac{\Gamma'(E)}{\Gamma(E)}(1-t)\mu(E)\right) + \frac{\Gamma'(E)}{\Gamma(E)}\left((1-t)\mu(E) + Y_\theta\right) + (t-s)\mu'(E) - s\beta^{\varphi}(E)\frac{\mu(E)}{E}\right]\delta E \\ &\quad + \lambda\mu(E)(1-\xi)\delta t,\end{aligned}$$

or, using $\delta t = -(1-t)\Delta\beta(E)\delta E/E$ and the definition of t_{Pigou}, $\delta W = \lambda\frac{\mu(E)}{E}\left[\xi(1-t) + (t-s) - t_{Pigou}\right]\delta E$ after some algebra. Since $\delta W = 0$ at the optimum, the result follows.

Part (ii). The inner problem given E and t solves (6) s.t. (7), (22) and (23). When $s = 0$, this is the same problem as without the tax t except for the fact that (i) wages w are after-t wages and (ii) there are the additional constants $t\mu(E)$ in (22) and (23). As a result, the optimal marginal tax formula (11) goes through, where w is the after-t wage and the effective marginal keep share on rent-seeking income is given by $(1-t)(1-T'(y(w)))$. Hence, the top marginal tax rate in the rent-seeking activity is

$$1-(1-t)(1-T'(y(\overline{w}_{E,t}))) = 1-(1-t)\left(1-\xi\frac{f^{\varphi}_{E,t}(\overline{w}_{E,t})}{f_{E,t}(\overline{w}_{E,t})}\right) = t + (t_{Pigou}-t)\frac{f^{\varphi}_{E,t}(\overline{w}_{E,t})}{f_{E,t}(\overline{w}_{E,t})},$$

where the last step uses the result from part (i).

Part (iii). To prove the result, we first show how the outer problem decomposition is extended to the presence of the linear tax t. Fix the rent-seeking tax t and $s = 0$. If $t = 0$, $\partial W(E,t)/\partial E$ is given by the right-hand-side of (13), since the outer problem for E is identical to the baseline without a rent-seeking tax.

[29] Note that this does not affect the multipliers λ and ξ: it is a change of variables within the integrals, not a reformulation of the constraints.

If $t \neq 0$, one can derive $\partial W(E,t)/\partial E$ using the same steps as in the proof of Lemma 3. The presence of t changes this derivation in two ways. First, the consistency and resource constraints (22) and (23) contain an extra $t\mu(E)$; this gives rise mechanically to an extra $(1-\xi)\lambda t\mu'(E)$ term. Second, wages w are after-t. This changes the interpretation (but not the form) of the effects R, I, C and S. The only formal change is when going from the first to the second line of (44), because now $\int_{\underline{w}_{E,t}}^{\overline{w}_{E,t}} we(w)f_{E,t}^{\varphi}(w)dw = (1-t)\mu(E)$. Hence, the last line of (44) becomes

$$-\lambda\frac{\mu(E)}{E}t_{Pigou} + \lambda\frac{\mu(E)}{E}\beta^{\varphi}(E)t + \xi\lambda(1-t)\beta^{\varphi}(E)\frac{\mu(E)}{E} + \xi\lambda\int_{\underline{w}_{E,t}}^{\overline{w}_{E,t}} e(w)(K_E(w)+L_E(w))dw.$$

Incorporating these two changes implies that, for any given (E,t),

$$\frac{\partial W(E,t)}{\partial E} = \lambda\frac{\mu(E)}{E}\left(\xi - t_{Pigou} + t(1-\xi)\right) + \frac{\Delta\beta(E)}{E}\left[I + R + \xi\lambda\left(C+S\right)\right]. \tag{54}$$

The result then follows from the fact that setting $t = t_{Pigou}$ yields $\partial W(E,t)/\partial E \equiv 0$ using (54) since $\xi = 0$ by the result in part (i) and $I = R = 0$ when there are no redistributive motives ($\eta(w) = 0$ and $\Psi(\theta,\varphi) = F(\theta,\varphi)$). Moreover, with $\eta(w) = 0$, $T'(y(w)) = 0$.

B Data and Identification

We use the March releases of the CPS for 2013 and 1979. We use the earnings data and the self-reported estimate of hours worked to construct wages. The CPS also provides a detailed industry classification for working individuals. We drop individuals for whom earnings, hours, age or industry is not reported. Following Heathcote et al. (2010), we also restrict attention to working age individuals between ages 25 and 65 and those employed (dropping those with very low hours or earnings per year). All variables are weighted with the provided weights and dollar denominated variables from 1979 are inflated to 2013 dollars using the CPI.

The industry variable we use to classify individuals, for the sake of our illustration, as rent-seekers or traditional workers is a 3-digit NAICS-based industry code. For 2013, this is the variable ind02 in the NBER released Merged Outgoing Rotation Groups (MORG) data. We choose the finance and law-related categories "Banking and related activities," "Savings institutions, including credit unions," "Non-depository credit and related activities," "Securities, commodities, funds, trusts, and other financial investments," "Insurance carriers and related activities," "Real estate," and "Legal services" to capture rent-seeking (admittedly somewhat arbitrary, but following the spirit of, and simplifying comparison to, Lockwood et al., 2013). For 1979, we use the comparable 3-digit variable ind70 with similar categories "Banking," "Credit agencies," "Security, commodity brokerage, and investment companies," "Insurance," "Real estate, incl. real estate-insurance-law offices," "Finance, insurance, and real estate–allocated," and "Legal services."

We use the sectoral wage data in both years to estimate an underlying bivariate lognormal wage distribution following Heckman and Honoré (1990) and French and Taber (2010). Let $(w_{\theta i}, w_{\varphi i})$ be individual i's potential log-wages in the traditional and rent-seeking activities, respectively, and assume

$$\begin{pmatrix} w_{\theta i} \\ w_{\varphi i} \end{pmatrix} = N\left(\begin{pmatrix} \mu_\theta \\ \mu_\varphi \end{pmatrix}, \begin{pmatrix} \sigma_\theta^2 & \sigma_{\theta\varphi} \\ \sigma_{\theta\varphi} & \sigma_\varphi^2 \end{pmatrix}\right).$$

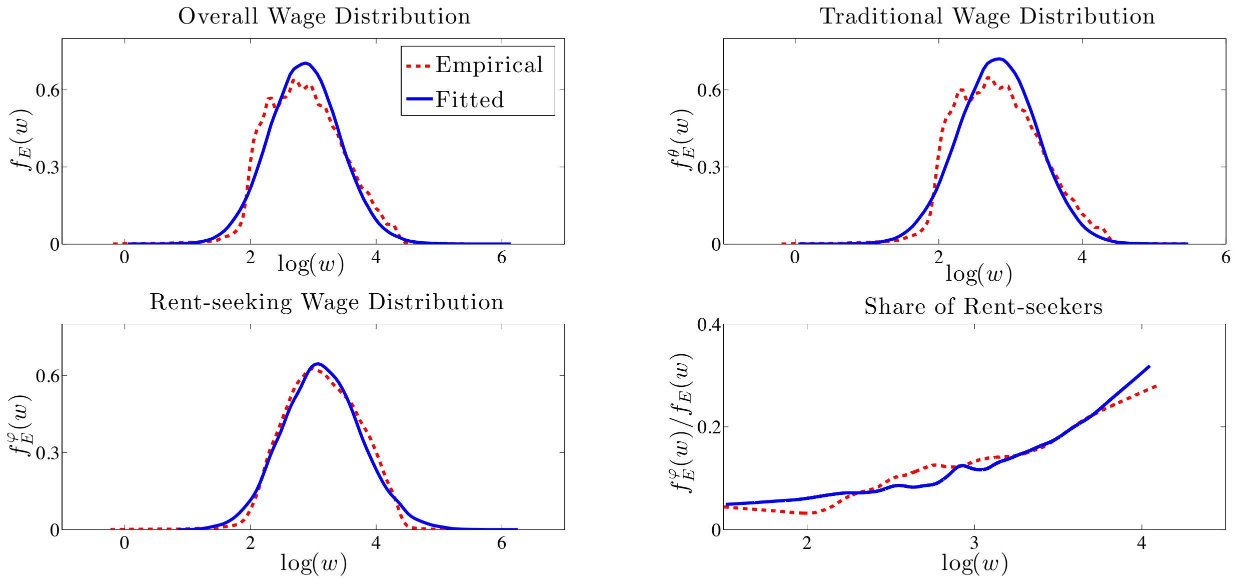

Figure 3: Empirical and fitted (sectoral) log-wage distributions

Letting $\lambda(.) = \phi(.)/\Phi(.)$ be Mill's ratio (and ϕ and Φ the pdf and cdf of a standard normal, respectively),

$$c = \frac{\mu_\theta - \mu_\varphi}{\sqrt{\sigma_\theta^2 + \sigma_\varphi^2 - 2\sigma_{\theta\varphi}}}$$

and

$$\tau_j = \frac{\sigma_j - \sigma_{\theta\varphi}}{\sqrt{\sigma_\theta^2 + \sigma_\varphi^2 - 2\sigma_{\theta\varphi}}}, \quad j = \theta, \varphi,$$

then we can estimate μ_j, σ_j^2 and $\sigma_{\theta\varphi}$ by matching the conditional moments

$$\begin{aligned}
Pr(J_i = \theta) &= \Phi(c) \\
\mathbb{E}[w_i | J_i = \theta] &= \mu_\theta + \tau_\theta \lambda(c) \\
\mathbb{E}[w_i | J_i = \varphi] &= \mu_\varphi + \tau_\varphi \lambda(-c) \\
Var[w_i | J_i = \theta] &= \sigma_\theta^2 + \tau_\theta^2(-\lambda(c)c - \lambda^2(c)) \\
Var[w_i | J_i = \varphi] &= \sigma_\varphi^2 + \tau_\varphi^2(\lambda(-c)c - \lambda^2(-c)),
\end{aligned}$$

where w_i and J_i are individual i's observed wage and sector, respectively. Using this procedure, we obtain $\mu_\theta = 2.8201, \mu_\varphi = 2.0056, \sigma_\theta^2 = 0.5488, \sigma_\varphi^2 = 0.8786, \rho_{\theta\varphi} \equiv \sigma_{\theta\varphi}/(\sigma_\theta \sigma_\varphi) = 0.5943$ for 2013. Figure 3 plots the resulting distributions of log-wages (both overall and for the two sectors) as well as the share of rent-seekers as a function of the wage, both for the data and the fitted bivariate lognormal, which demonstrates a reasonably good fit (except at the very top, where the wage distribution is well known not to be lognormal but Pareto in the data). We truncate the resulting skill distribution at the top 0.1 percentile in both dimensions and rescale accordingly.

To compute optimal income taxes, we begin with the inner problem for given E. The (sectoral) wage

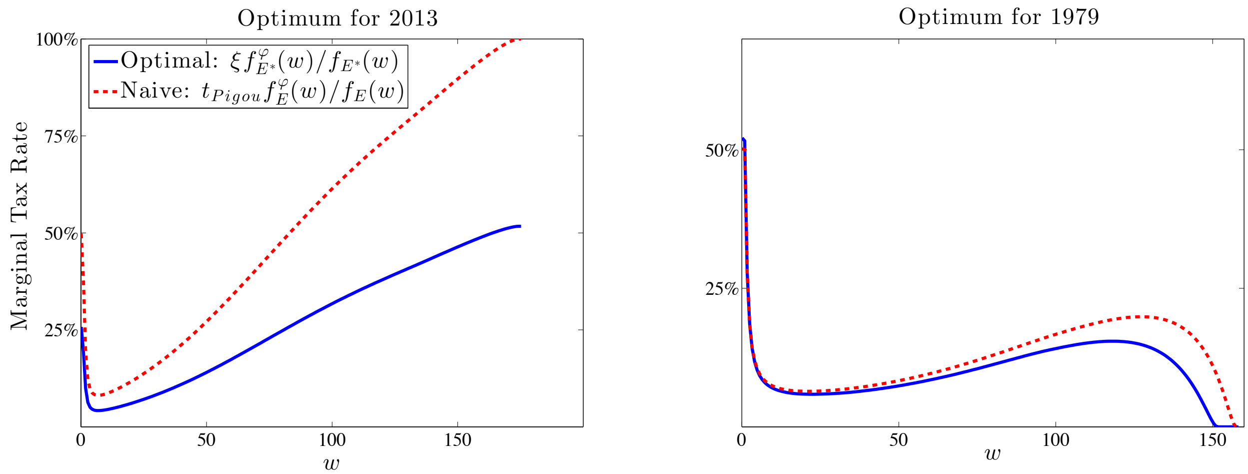

Figure 4: Optimal policy for 2014 versus 1979

distributions are obtained numerically, and the resulting optimal tax problem can be solved using the same methods as for a standard Mirrlees model (making use of the local incentive constraints), with the only additional complication that the multiplier ξ needs to be found numerically such that the consistency constraint is satisfied. We then repeat this procedure to find the optimal E using a grid search. We finally verify that the monotonicity constraint $y'(w) \geq 0$ is satisfied, so the solution is globally incentive compatible.

We can perform the same steps based on the 1979 CPS data. This allows us to compare the optimal policies for these periods, as alluded to in Section 5. As can be seen from Figure 4, rent-seeking was less important at the top of the skill distribution in 1979, so that the share of rent-seekers converges to 0 for high wages given the calibrated skill distribution, in contrast to 2013. Moreover, whereas the optimal correction ξ is .53 for 2013, it increases to .92 for 1979. These comparisons suggest that, while the divergence between the optimal correction ξ and the Pigouvian correction $t_{Pigou} = 1$ was not as pronounced in the 1970s, it has become much more important today. Roughly half of this difference is due to changes in the underlying skill distribution, whereas the other half is due to the higher level of rent-seeking according to our definition (recall $\xi = .7$ for the 2013 skill distribution with 1979 rent-seeking levels).