FUNDAÇÃO GETULIO VARGAS

EPGE

Escola de Pós-Graduação em Economia

Ensaios Econômicos

Escola de Pós-Graduação em Economia da Fundação Getulio Vargas

N° 775 ISSN 0104-8910

Optimal Mirrleesian Taxation in Non-competitive Labor Markets

Carlos da Costa, Lucas Maestri

Dezembro de 2015

URL: http://hdl.handle.net/10438/15026

da Costa, Carlos
Optimal Mirrleesian Taxation in Non-competitive Labor Markets/ Carlos da Costa, Lucas Maestri - Rio de Janeiro : FGV,EPGE, 2015
34p. - (Ensaios Econômicos; 775)

Inclui bibliografia.

CDD-330

Optimal Mirrleesian Taxation in Non-competitive Labor Markets

Carlos da Costa　　Lucas Maestri

December, 2015

Abstract

We study optimal labor income taxation in non-competitive labor markets. Firms offer screening contracts to workers who have private information about their productivity. A planner endowed with a Paretian social welfare function tries to induce allocations that maximize its objective. We provide necessary and sufficient conditions for implementation of constrained efficient allocations using tax schedules. All allocations that are implementable by a tax schedule display negative marginal tax rates for almost all workers. Not all allocations that are implementable in a competitive setting are implementable in this non-competitive environment. ***Keywords:*** *Mirrlees' problem; Non-competitive labor markets.* ***JEL Classification:*** *D82, H21.*

Optimal tax theory in the Mirrlees' (1971) tradition implicitly relies on a view of perfectly competitive labor markets. Under perfect competition a worker faces several firms offering him employment at a wage equal to his marginal productivity, when he is searching for a job. A worker who quit his job always finds a similar one at the same wage. This is hardly how economists currently view the job market: workers usually cherish having a good job and are desolated when they lose one. Accordingly, economists widely acknowledge the distance between the perfect competition paradigm and way markets function in the real world. see Manning (2005); Reich (2015).

Lack of competition is not only present in many accounts of the actual workings of labor markets but this very idea underlies many actual policies, most prominent of which is the minimum wage. The sheer presence of market power potentially leads to inefficiencies and/or inequities which can be alleviated with government intervention. However, in contrast with the literature which addressed labor market policies under competitive markets, very little is known about constrained efficient policies in this environment. Whereas the latter has evolved by adopting a mechanism design approach, the former is almost exclusively done by restricting policy instruments and asking how to optimally use them.[1]

In this paper we take a step towards bringing the analysis of optimal tax policy under imperfect competition to the same footing as that under competitive markets. We address optimal tax policy in a non-competitive labor market using a mechanism design approach. We modify Mirrlees' (1971) canonical model to accommodate the presence of market power. Agents are heterogeneous with respect to the utility cost

[1] A noteworthy recent exception is Cahuc and Laroque (2014).

of producing any given amount of output and this is private information. Each firm is randomly matched with workers whose productivities are private information. The firm offers screening contract to workers. This protocol under which firms move first making take it or leave it offers generates substantial market power to firms.[2]

With market power by firms is that earnings need not be equal to the output an worker produces. Another layer of informational asymmetry between workers and the government is introduced since the latter can only observe earnings. [3] This raises new issues as government goal is to design tax policies which are optimally used as screening instruments for a firm which itself is trying to screen a worker who has private information on his productivity. Therefore the usual approach used in the screening literature — e.g. Mirrlees (1971) and Baron and Myerson (1982) — does not apply.

To circumvent this difficulty we develop a new approach based on a duality between the planner's and the firm's Euler equation. It is valid under the assumption that the taxation principle applies in our setting, i.e., that all that may be implemented via mechanisms may also be implemented through budget sets and vice-versa.[4] Indeed, by transforming both the firm's and the planner's programs into calculus of variation problems we are able to derive necessary and sufficient conditions for implementation via taxes. We then characterize optimal tax schedules for which the necessary conditions for implementation are satisfied.

Tax schedules must naturally be based on variables that are observed by the Government. This requirement is satisfied by tax systems comprised of labor income tax schedules, taxes on firms' profits, and transfers to unemployed workers in the form of unemployment benefits. We show that for an allocation which maximizes a utilitarian social welfare function to be induced by a tax schedule, marginal tax rates must be negative for all but the workers at the top of the distribution. Such schedule must be supplemented by a cash transfer for any worker who does not work, and by a tax on the firm's profits.[5]

To understand the rationale for marginal earnings subsidies, start from the optimal utilitarian allocation. Then, consider the impact of reducing labor supply of a worker with the highest possible disutility of effort. To keep this agent's utility constant, this must be accompanied by lower earnings for the worker. Such a reform makes the bundle less attractive for all workers with lower disutility of effort, i.e., it lowers their utility if they deviate and choose this bundle. Such a reform is attractive to the firm since the utilitarian allocations leaves too much rent in the hands of all workers from the firm's perspective. To implement this allocation the firm must distort effort downwards. Finally, to counter the distortion introduced

[2]It is not hard to see that each firms is in effect a monopsonist with regards to the workers with whom it is matched. Of course market power may also be a consequence of there simply being few firms operating in the market — see Manning (2005); Reich (2015). We compare the consequences of adopting this view for the source of market power in Section 5.

[3]Our setting is in this sense related to Stantcheva's (2014). There, workers of different productivities apply to work for firms which are able to observe their effort but not their output. A non-linear schedule specifies the pay for a given number of hours. As in our model, the planner who observes neither effort (hours) nor output must base its policy on earnings.

[4]See Hammond (1979); Guesnerie (1998).

[5]To rule out trivial implementation, i.e., those attained by making the firm indifferent between all possible allocations, we assume that net profits are strictly increasing on gross profits, although we may be arbitrarily close to 100%.

by the firm the planner imposes a negative marginal tax rate on workers' earnings. Although we have focused on the worker with the highest disutility of effort, the same is true for all other workers with the proviso that the reform must also take into account incentive effects for workers with higher disutility of effort.

The characterization results mentioned above are based on the necessary conditions for implementation. It is not always the case that an allocations that solves the planner's utilitarian program can be implemented when labor markets are not competitive. Indeed, we offer an example of a Pareto efficient allocation for which the necessary conditions for implementation fail to hold.

When, however, the sufficient conditions are also satisfied it is possible to show that tax schedules are progressive. Marginal tax rates are non-positive but increasing in earnings.

We then generalize the results above by considering different Pareto weights for the planner's program. More precisely, we consider a sequence of distribution of weights that keep increasing the relative weights of the most disadvantaged workers. Strictly negative marginal tax rates obtain for any worker for whom we there are (positive measure of) workers with lower disutility of effort and strictly positive Pareto weights. We show that as the planner's metric approaches a Rawlsian limit the sufficient conditions for implementation are eventually satisfied. For every agent, the marginal tax rates becomes arbitrarily close to zero. In the limit case for which the planner wants to implement a Rawlsian allocation, marginal tax rates are zero for all workers and the optimal policy is comprised of a transfer to the unemployed financed by taxes on profits.

Thus far, we have assumed that all that can be implemented through game forms can also be implemented via budget sets. But is this the case? As we have mentioned, firms have in practice, monopsonistic power over the workers they hire. In this sense, any implementation that is possible under the market structure we examine is also feasible under a monopsony. The converse is not true, however. We show that monopsonist markets pose no impediments to constrained efficiency *if* the planner has access to very rich information about the firm's payroll. Any (constrained) efficient allocation may be implemented by harshly punishing the firm if it does not generate the distribution of earnings that is associated with the desired allocation. Note that such mechanism requires the planner to use the information on all the distribution of earnings and the fact that all workers in the economy work for the same firm. Such a mechanism is, therefore, quite sensitive to the details of the environment and therefore unlikely to be useful for police prescriptions.

The market structure we consider is the polar opposite view from the competitive paradigm used in the literature. By focusing on this benchmark case, we could enhance our comprehension of the interaction between imperfect competition in the labor market and optimal redistributive policies. Accordingly, several economic forces identified by our model are likely to emerge in models characterized by some, even if not monopsonistic, power.

The rest of the paper is organized as follows. After this introduction, Section 1 describes the environment. Section 2 is where we describe the nature of the problem we study. Section 2.1 describes the program which defines the allocation that the planner wishes to implement. In Section 2.2 we show the problem faced by firms when given the tax schedule imposed by the Government. The main results of the paper are in Section 3. In Section 4 we extend the planner's objective beyond utilitarianism. We discuss alternative mechanisms for which the direct reliance on crossing information between workers and firms are crucial in Section 5. Section 6 concludes. Very technical derivations are relegated to appendices.

1 Environment

We consider an economy inhabited by a continuum of workers that differ from one another according to a single preference parameter θ, that defines the dis-utility of effort. Formally, workers have preferences defined over consumption, $c \in \mathbb{R}_+$, and effort, $n \in \mathbb{R}_+$, of the following form.

$$U(c, n, \theta) = \nu(c) - \theta\eta(n).$$

We assume that ν and η are smooth, satisfy $\nu', -\nu'', \eta' > 0$ and $\eta'' \geq 0$. We normalize $\nu(0)$ and $\eta(0)$ to 0. The parameter θ belongs to an interval $\left[\underline{\theta}, \overline{\theta}\right] \subset \mathbb{R}_{++}$ and captures the disutility from working. Under separable iso-elastic preferences the model is isomorphic to Mirrlees' (1971) optimal tax setting, where it is productivity that differs across workers. As in Mirrlees (1971), we assume that θ is workers' private information.

The distribution of types, $F(\cdot)$, which is common knowledge has an associated smooth density $f(\cdot)$.

Thus far our model is exactly Mirrlees's (1971) specialized to separable preferences. We now depart from the competitive assumption that underlies Mirrlees' (1971) and, for that matter, almost all the optimal tax literature. We consider a continuum of workers and firms, in which each firm is randomly matched with a finite number or workers and every worker is matched with a single firm.

The timing of the model is as follows. First, the government sets a lump-sum tax on all firms which decide to participate in the market. Upon agreeing to participate, the firm is randomly matched with workers each of whom has private information about its productivity. Then the firm offers a screening contract to workers with whom it is matched.

Firms' ownership is spread uniformly across workers in such a way that no worker wishes the firm to deviate from profit maximizing behavior. A firm operates as long as its expected profits are non-negative. It is apparent that our model is isomorphic to one for which a single firm operates in the market. We discuss this alternative modelling assumption in Section 4.

In Mirrlees' (1971) formulation of the problem, the source of informational asymmetry between the planner and the workers is the fact that effort is not observed by the planner, only output. Because the latter is the product of a worker's productivity, which is only privately observed, and his effort, the observation of output does not allow the planner to disentangle the former from the latter. That is, output is observed but the utility cost of producing such output is not observed. Here, the planner observes neither workers' effort nor the output generated by such effort. All it observes are the earnings of each worker. It must therefore base policy on earnings, which need not coincide with output in this non-competitive setting.

For our purposes it will be convenient to define the inverse functions $\phi = \eta^{-1}$ and $\kappa = \nu^{-1}$, and work directly with the variables h and u defined by $\nu(\kappa(u)) = u$ and $\eta(\phi(h)) = h$.

2 Setting Up the Stage

2.1 The Mirrlees' (1971) Program

Assume that the planner could observe workers' outputs. The planner would then choose an allocation that solved the Utilitarian program — henceforth, Program $\mathcal{P}$,[6]

$$\max \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h(\theta)) - \kappa(u(\theta))\right] f(\theta) d\theta \tag{1}$$

subject to,

$$v(\theta) = v(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} h(\tilde{\theta}) d\tilde{\theta}, \tag{2}$$

$$v(\theta) = u(\theta) - \theta h(\theta) \tag{3}$$

$$h : \theta \longmapsto \mathbb{R}_{+}, \text{non-increasing.} \tag{4}$$

and

$$\int_{\underline{\theta}}^{\bar{\theta}} v(\theta) f(\theta) d\theta \geq w. \tag{5}$$

We make assumptions about the nature of the Mirrlees' (1971) program's solution. That is, we focus on problems which primitives lead to well behaved solutions for program $\mathcal{P}$ that display no exclusion and no bunching. For simplicity, we also restrict our analysis to cases for which the solution is characterized by smooth consumption and labor-supply profiles.

Assumption 1. *Let $(v^{*}(\theta))_{\theta \in [\underline{\theta}, \bar{\theta}]}$ be the solution to the Mirrlees' (1971) Problem, $\mathcal{P}$. Let $(h^{*}(\theta))_{\theta \in [\underline{\theta}, \bar{\theta}]}$ be the labor supply induced by v^{*} and let $(u^{*}(\theta))_{\theta \in [\underline{\theta}, \bar{\theta}]}$, the profile of utilities from consumption induced by v^{*}. Then,*

(i) $v^{*} : \theta \rightarrow v^{*}(\theta)$ is smooth, and;

(ii) $\dot{v}^{*}$ is strictly increasing.[7]

For any incentive-compatible and differentiable profile of indirect utilities, $(v(\theta))_{\theta \in [\underline{\theta}, \bar{\theta}]}$, the amount of labor supplied by type θ, $h(\theta)$, satisfies $\dot{v}(\theta) = -h(\theta)$ and his flow utility from consumption is recovered from $v(\theta)$ and $\dot{v}(\theta)$ through $u(\theta) = v(\theta) - \theta \dot{v}(\theta)$.

Our analysis is based on the relationship between the planner's and the firm's problem. We review here the some basic properties of $\mathcal{P}$'s solution. Under Assumption 1, the planner's program can be written by adding a multiplier, λ, to the isoperimetric constraint (5) — see Theorem 14.21 in Clarke (2013). We may therefore write $\mathcal{P}$, as

$$\max \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi\left(-\dot{v}(\theta)\right) - \kappa\left(v(\theta) - \theta \dot{v}(\theta)\right) + \lambda v(\theta)\right] f(\theta) d\theta - \lambda w.$$

[6] In Section 4 we consider alternative Paretian objectives for the planner.

[7] We use a dot over a function to denote its derivative with respect to θ, e.g. $\dot{v}(\theta) = dv(\theta)/d\theta$.

The solution $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ of $\mathcal{P}$ satisfies the Euler equation,

$$\frac{d}{d\theta}\left\{\left[\phi'(-\dot{v}^*(\theta)) - \theta\kappa'(v^*(\theta) - \theta\dot{v}^*(\theta))\right]f(\theta)\right\} = \left[\kappa'(v^*(\theta) - \theta\dot{v}^*(\theta)) - \lambda\right]f(\theta). \quad (6)$$

Integrating both sides of (6) and using the fact that the allocation of the most efficient type, $\underline{\theta}$, is not distorted, $\phi'(-\dot{v}^*(\underline{\theta})) = \underline{\theta}\kappa'(v^*(\underline{\theta}) - \underline{\theta}\dot{v}^*(\underline{\theta}))$, we get

$$\left[\phi'(-\dot{v}^*(\theta)) - \theta\kappa'(v^*(\theta) - \theta\dot{v}^*(\theta))\right]f(\theta) = \int_{\underline{\theta}}^{\theta}\left[\kappa'(v^*(a) - a\dot{v}^*(a)) - \lambda\right]f(a)da. \quad (7)$$

In all that follows $(h^*(\theta), u^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ will denote the allocation associated with $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$.

Is it possible to implement $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$, in our setting? I.e., is there a tax system that induces the planner's allocation as an equilibrium for our economy?

It will soon be clear that a policy comprised of a single labor income tax schedule will fail in general to implement the planner's allocation. The reason is twofold. First, the firm will always set the utility of the least productive type that it hires as low as possible – lemma 1, below. Therefore, to guarantee that the lowest type that is hired by the firm obtains utility $v^*(\bar{\theta})$ the government will have to pay unemployment benefits, b, which deliver utility $\nu(b) = v^*(\bar{\theta})$ for anyone who is not working.

Second, because the policy entails transfers to workers that increase firms' profits, the planner will use taxes on profits to balance the budget. However, to rule out a trivial implementation in which the government tax 100% of firms' profits and makes firms indifferent between all allocations, we consider taxes such that net profits are strictly increasing in gross profits, albeit arbitrarily close to 100%.

2.2 The Firm's (relaxed) Program

Assume that the planner has imposed an earnings tax schedule, $\mathcal{T}(\cdot)$, i.e., a mapping from earnings into tax obligations.

We define $\chi(\cdot)$ as the function that maps the flow utility from consumption that the firm delivers to the worker to total earnings that the worker needs in order to attain this level of flow utility. Provided that χ is a monotonic function it is related to $\mathcal{T}$ through $\nu(\chi(u) - \mathcal{T}(\chi(u))) = u$. To account for possible non-monotonicity of χ we shall refer to the taxes paid by an agent with flow utility of consumption u as the value defined through $\chi(u) - T(u) = \kappa(u)\ \forall u \in \mathbb{U}$.

Assume that $\chi : \mathbb{U} :\to \mathbb{R}$ is continuously differentiable over this domain. This will allow us to use standard optimal control methods.

Each worker's consumption is then his earnings minus the taxes he must pay. Let $u(\theta)$ be the flow utility from consumption that a type θ worker obtains and $\mathbb{U}$, the set of flow utilities from consumption that the firm at which he works can deliver.[8] We let $\mathbb{U} := [u_1, u_2]$ be a compact interval of $\mathbb{R}_{++}$ containing $\left[u^*(\bar{\theta}), u^*(\underline{\theta})\right]$. To simplify the analysis below we assume that $u_1 < u^*(\bar{\theta})$ and $u_2 > u^*(\underline{\theta})$. Simplicity stems

[8]Since c is observable and u is strictly monotonic, $u = \nu(c)$ can be recovered from c.

from the fact that under this assumption the flow utility from consumption, $u(\theta)$, of any worker, θ, is an interior point of $\mathbb{U}$ along the Mirrlees's (1971) solution profile, $(u^*(\theta))_{\theta\in\Theta}$. Beyond simplicity, this assumption guarantees that the implementation we consider does not hinge upon any discontinuity of the tax schedule around the desired allocation.

The firms' relaxed program (Program $\mathcal{P}^F$) is then

$$\max \int_{\underline{\theta}}^{\theta_0} \left[\phi(h(\theta)) - \chi(u(\theta))\right] f(\theta)d\theta$$

s.t., (2), (3), and

$$v(\theta_0) \geq v_0,$$

where $\theta_0 \leq \bar{\theta}$ is the highest type potentially hired by the firm and v_0 is the utility that workers obtain if they are not working. Note that we have not imposed the monotonicity constraint (4) in Program $\mathcal{P}^F$'s definition.

The first thing we note is that, if a solution to Program $\mathcal{P}^F$ exists, then it leaves no rent to the least efficient type that is hired — lemma 1.

Lemma 1. *If a solution to the firm's relaxed program exists then it delivers utility* v_0 *to the least efficient type that is hired.*

Proof. Let v be a solution to the firm's problem and let $\hat{\theta}$ be the highest type hired. Assume towards a contradiction that $v(\hat{\theta}) > v_0$. Incentive compatibility implies that $\hat{\theta} = \bar{\theta}$. Let, then, $\{h, u\}$ be the functions associated with v. Take $\varepsilon > 0$ such that $u(\bar{\theta}) - \bar{\theta}(h(\bar{\theta}) + \varepsilon) > v_0$.

Notice that the new allocation $(\check{h}, \check{u})$ defined by $\check{h}(\theta) := h(\theta) + \varepsilon$ for all θ and $\check{u}(\theta) := u(\theta)$ for all $\theta \leq \hat{\theta}$ is implementable and yields a strictly higher profit for the firm. A contradiction. □

The intuition behind this result is straightforward. If v solves the firm's problem and the highest type that is hired obtains a positive rent, then the firm can implement a new allocation in which all hired types have the same consumption level, but all of them supply more labor.

Assume that the tax policy is such that the firm chooses to hire any type, i.e., $\theta_0 = \bar{\theta}$.[9] Then, using the same arguments used in Section 2.1, we can write Program $\mathcal{P}^F$ as the following calculus of variation problem,

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(-\dot{v}(\theta)) - \chi(v(\theta) - \theta\dot{v}(\theta))\right] f(\theta)d\theta$$

s.t.,

$$v(\bar{\theta}) \geq v_0.$$

3 Implementing the Mirrlees' (1971) allocation

Assume that a function $\chi(\cdot)$ exists for which the firms' program $\mathcal{P}^F$ exists and coincides with the Mirrlees' (1971) solution, $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$.

[9] In sections 3.1 and 3.2 we offer necessary and sufficient conditions on the level of taxes for this to be the case.

In Section 3.1 we characterize such a function, assuming that it exists. In Section 3.2 we show conditions under which it does exist. In particular, we derive sufficient conditions which depend only on the properties of the distribution of types. We also allow for bunching in Section 3.3.

3.1 Characterizing Optimal Tax Systems

If the utility profile, $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$, that solves the Mirrlees' (1971) also solves the firm's relaxes program $\mathcal{P}_1$, then it must satisfy the Euler equation,

$$\frac{d}{d\theta}\left\{\left[\phi'\left(-\dot{v}^*(\theta)\right)-\theta\chi'(v^*(\theta)-\theta\dot{v}^*(\theta))\right]f(\theta)\right\}=\chi'(v^*(\theta)-\theta\dot{v}^*(\theta))f(\theta). \tag{8}$$

Since $u^*(\underline{\theta})<u_2$, standard arguments imply that the most efficient worker is not distorted, i.e., $\phi'(-\dot{v}^*(\underline{\theta}))=\underline{\theta}\chi'(v^*(\underline{\theta})-\underline{\theta}\dot{v}^*(\underline{\theta}))$. Since there is no distortion in the solution of the planner's problem, $\kappa'(v^*(\underline{\theta})-\underline{\theta}\dot{v}^*(\underline{\theta}))=\phi'(-\dot{v}^*(\underline{\theta}))$, we conclude that the most efficient worker faces zero marginal tax rates: $T'^*(\underline{\theta})))=0$.

Integrating (8) we obtain

$$\left[\phi'\left(-\dot{v}^*(\theta)\right)-\theta\chi'(v^*(\theta)-\theta\dot{v}^*(\theta))\right]f(\theta)=\int_{\underline{\theta}}^{\theta}\left[\chi'(v^*(a)-a\dot{v}^*(a))\right]f(a)da. \tag{9}$$

We have reached these conclusions ignoring the monotonicity constraints in the firm's program, $\mathcal{P}^F$. Recall, however, that we have assumed that the solution for the Mirrlees' (1971) program, $\mathcal{P}$, is monotonic. Hence these constraints do not bind in the sense that a small perturbation of the Mirrlees' (1971) solution $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ satisfies them. The Euler equation (8) and the transversality condition,

$$\phi'(-\dot{v}^*(\underline{\theta}))=\underline{\theta}\chi'(v^*(\underline{\theta})-\underline{\theta}\dot{v}^*(\underline{\theta})),$$

must therefore be satisfied.

Lemma 2. *If $\chi(\cdot)$ implements the Mirrlees' (1971) solution, $v^*(\theta)$, then (9) holds at the solution to the firm's problem and the marginal income tax on the most efficient worker is zero.*

To further characterize the the structure of a tax system which implements the allocation $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ we evaluate (9) and (7) at the Mirrlees' (1971) allocation and ask which properties χ must possess to induce v^*. This procedure along with Lemma 2 yield the following proposition.

Proposition 1. *If $\chi(\cdot)$ implements the Mirrlees' (1971) solution, $v^*(\theta)$, then*

$$\chi'(u^*(\theta))=\kappa'(u^*(\theta))-\frac{\lambda}{f(\theta)\theta^2}\int_{\underline{\theta}}^{\theta}f(\tilde{\theta})\tilde{\theta}d\tilde{\theta},\ \forall\theta.$$

I.e., marginal tax rates are negative for all workers $\theta>\underline{\theta}$.

Proof. We combine the planner and the firm's Euler equations by subtracting (9) from (7), to obtain

$$\theta\left[\kappa'(u^*(\theta)) - \chi'(u^*(\theta))\right] f(\theta) = -\int_{\underline{\theta}}^{\theta} \left[\kappa'(u^*(a)) - \chi'(u^*(a)) - \lambda\right] f(a)da. \tag{10}$$

Let us define

$$\rho(\theta) \equiv \left[\kappa'(u^*(\theta)) - \chi'(u^*(\theta))\right] f(\theta),$$

and re-write (10) as

$$\theta\rho(\theta) = -\int_{\underline{\theta}}^{\theta} \left[\rho(a) - \lambda f(a)\right] da. \tag{11}$$

Differentiating (11) with respect to θ and multiplying by θ we get

$$\frac{d}{d\theta}\left(\rho(\theta)\theta^2\right) = \lambda f(\theta)\theta$$

Integrating this expression, yields

$$\rho(\theta)\theta^2 = \rho(\underline{\theta})\underline{\theta}^2 + \lambda \int_{\underline{\theta}}^{\theta} f(\tilde{\theta})\tilde{\theta}d\tilde{\theta}.$$

Since we have no distortion at the bottom for both the planner's and the firm's problem, we have $\rho(\underline{\theta}) = 0$. We may then write

$$\rho(\theta) = \frac{\lambda}{\theta^2}\int_{\underline{\theta}}^{\theta} f(\tilde{\theta})\tilde{\theta}d\tilde{\theta} \geq 0 \tag{12}$$

Recall now that

$$-\frac{\rho(\theta)}{f(\theta)} + \kappa'(u^*(\theta)) = \chi'(u^*(\theta)),$$

which implies

$$\chi'(u^*(\theta)) = \kappa'(u^*(\theta)) - \frac{\lambda}{f(\theta)\theta^2}\int_{\underline{\theta}}^{\theta} f(\tilde{\theta})\tilde{\theta}d\tilde{\theta}. \tag{13}$$

We say that a tax system involves marginal incentives to consumption if the consumer is being subsidized at the margin: $\chi'(u) < \kappa'(u)$. □

When the firm contemplates increasing the labor supply of some type θ, it trades-off the efficiency gains from this policy with the losses associated with leaving larger information rents to every type $\tilde{\theta} < \theta$, which the firm does not value. The government, on the other hand, benefits directly from the additional rent left to types $\tilde{\theta} < \theta$, since it places positive weight on welfare of all workers. Thus, to align the firm's interests with its own interests the planner must lower the marginal cost of increasing labor supply of type θ. This is done by introducing a negative marginal tax rate.

The firm may also choose to exclude some types. This choice becomes very tempting when the firm makes small profits from workers with high disutilily of effort. Indeed, if the firm is not making profits with a worker it benefits from excluding him and increasing rents extracted from more efficient types. The firm would never, then, hire a worker and not make strictly positive profits on him. To guarantee that the firm hires all workers the planner must make sure that this is the case by subsidizing them, the intended allocation entails a level of consumption for the worker that exceeds his output.

Proposition 2. *Assume that the Mirrlees' (1971) allocation is such that the least efficient worker, $\bar{\theta}$, consumes no less than his output. Then a subsidy not inferior to*

$$\frac{1}{f\left(\bar{\theta}\right)}\int_{\underline{\theta}}^{\bar{\theta}}\phi'(h^*(\theta))\left(\frac{h^*\left(\bar{\theta}\right)}{\bar{\theta}}\right)f(\theta)d\theta,$$

must be given to all workers.

3.1.1 Wedges and Taxes

We have seen that marginal tax rates are always negative at the Mirrlees' (1971) allocation. This is in sharp contrast with Seade's (1977) findings regarding marginal tax rates in the Mirrlees' (1971) program, despite the fact that the same allocation is implemented.

To understand the difference, note that there are two layers of distortions which are imposed on each worker: one by the firm and the other by the planner. Wedges defined as the difference between the marginal productivity of labor and the marginal rate of substitution between consumption and effort are not, in general, equal to marginal tax rates. Indeed, let $\tau(\theta)$ denote the wedge faced by a type θ worker at the induced allocation. Then, by definition,

$$\tau(\theta)=\theta\frac{\kappa'(u(\theta))}{\phi'(h(\theta))}.$$

Next, let us consider the case in which $\chi(\cdot)$ is invertible, $\chi'(u(\theta))>0\ \forall\theta$. $\mathcal{T}(\chi)$ is defined in $[u^*(\bar{\theta}),u^*(\underline{\theta})]$ by $\mathcal{T}(\chi)=T\circ\chi^{-1}$. Thus the tax τ is obtained by: $\chi(u)=\kappa(u)+\tau(\chi(u))$. Of course, wedges and taxes are related. If $z(n)$ is the earnings function, mapping a workers' output, n, to his earnings, z, then, with some abuse in notation, $c(n)=z(n)-\mathcal{T}(z(n))$. This allows us to write $\tau=(1-\mathcal{T}')\,z'$. The earnings function, $z(n)$, is such that

$$z'(\phi(h(\theta)))=\theta\frac{\chi'(u(\theta))}{\phi'(h(\theta))},$$

for all θ.

Figure 1 displays marginal taxes, $\mathcal{T}'$, and wedges, τ, for example 3.2, below.

This is all based on the assumption that χ is strictly increasing in u along the whole interval $[u^*(\bar{\theta}),u^*(\underline{\theta})]$. We cannot however rule out the possibility that χ is decreasing in some interval, in which case the firm receives a transfer from the Government that exceeds the payment made to the worker. Due to the convexity of χ, we may find θ_a such that for all agents in the interval $[\bar{\theta},\theta_a]$ a firm pays no more than what it pays to $\bar{\theta}$ yet agents in this interval have higher earnings (and produce more).

Finally note that taxes, defined as the difference between earnings, z, and consumption, c, may differ from the total surplus generated by the worker, $n-c$, in this non-competitive environment . Indeed taxes are the sum of total surplus generated by the worker and the net profit the firm makes with the worker, $z-n$.

We can therefore read Proposition 3 as stating that whenever the desired allocation entails a negative surplus by the least efficient agent, the planner must offer a subsidy, $z-n<0$, to all workers.

3.2 Implementability

Is it always possible to implement the planner's solution?

Assume for the moment that the firm chooses to exclude no type. In this case, if the firm's program is concave in $(v, \dot{v})$, if v^* satisfies the Euler equation it solves the firm's program (see Theorem 18.8 in Clarke (2013)). It is not hard to see that the program is concave if and only if $\chi(\cdot)$ is a convex function.

Lemma 3. *Let $\mathbb{U} := [u_1, u_2]$ be a compact interval of $\mathbb{R}_{++}$ containing $\left[u^*(\bar{\theta}), u^*(\underline{\theta})\right]$. Assume that $\chi : \mathbb{U} :\to \mathbb{R}$ is continuous over this domain and that the firm's relaxed program $\mathcal{P}^F$ is concave in $\dot{v}$. If the firm chooses to exclude no type, then Program $\mathcal{P}^F$ has a solution.*

If χ induces the firm to hire all types then program $\mathcal{P}^F$ is equivalent to the calculus of variation problem above, which has a solution if it is concave in $\dot{v}$. In Proposition 3 we derive subsidies that guarantee that the firm will hire all types at the solution for the Mirrlees' (1971) Program.[10]

Differentiating (13) with respect to θ we get

$$-\frac{d}{d\theta}\left\{\frac{\lambda}{f(\theta)\theta^2}\int_{\underline{\theta}}^{\theta} f(\tilde{\theta})\tilde{\theta}d\tilde{\theta}\right\} + \kappa''(u^*(\theta))\dot{u}^*(\theta) = \chi''(u^*(\theta))\dot{u}^*(\theta), \tag{14}$$

where $\dot{u}^*(\theta) := \frac{d}{d\theta}u^*(\theta)$. By assumption, $\kappa''(\cdot) > 0$, whereas $\dot{u}^*(\theta) < 0$. Thus, a sufficient condition for $\chi'' > 0$ is

$$\frac{d}{d\theta}\left\{\frac{1}{f(\theta)\theta^2}\int_{\underline{\theta}}^{\theta} f(\tilde{\theta})\tilde{\theta}d\tilde{\theta}\right\} \geq 0. \tag{15}$$

Under (15) we can, in this case, construct a convex function $\chi(\cdot)$ from (13) and the boundary condition $\phi'(-\dot{v}^*(\underline{\theta})) = \underline{\theta}\chi'(v^*(\underline{\theta}) - \underline{\theta}\dot{v}^*(\underline{\theta}))$. This guarantees that the firm's problem is concave and that, provided that there is no exclusion, the Mirrlees' (1971) allocation, $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$, solves the firm's problem, $\mathcal{P}^F$

The conditions above do not pin down, $\phi\left(h^*(\bar{\theta})\right) - \chi(u^*(\bar{\theta}))$, hence the subsidy for the least efficient type, $\chi(u^*(\bar{\theta})) - \kappa(u^*(\bar{\theta}))$. Proposition 3, below, shows that we can construct a tax system in which the firm never excludes any type by offering to the least efficient worker, $\bar{\theta}$, a subsidy

$$\chi(\phi\left(h^*(\bar{\theta})\right) - \chi(u^*(\bar{\theta})) \geq \left[\phi\left(h^*(\bar{\theta})\right) - \phi\left(\frac{h^*(\bar{\theta})}{2}\right)\right] + \phi'\left(\frac{h^*(\bar{\theta})}{2}\right)\frac{h^*(\bar{\theta})}{\underline{f}}\left(\frac{\bar{\theta}}{\underline{\theta}^2}\right), \tag{16}$$

where $\underline{f} := \min_\theta f(\theta)$.[11]

Proposition 3. *Assume that (15) holds. Then, there is a tax system that implements the planner's problem. The mapping $\chi(\cdot)$ must satisfy (13), the boundary condition $\phi'(-\dot{v}^*(\underline{\theta})) = \underline{\theta}\chi'(v^*(\underline{\theta}) - \underline{\theta}\dot{v}^*(\underline{\theta}))$ and must give a subsidy (16) to the least efficient worker $\bar{\theta}$.*

[10] In fact we are able to provide sufficient conditions for the planner's allocation to solve the firm's profit maximizing problem, *even if the firm ignores the monotonicity constraints*. Hence, the solution to the relaxed program will solve the firm's profit maximization problem.

[11] A tighter bound,

$$\phi'\left(h^*(\bar{\theta})\right)\frac{h^*(\bar{\theta})}{\underline{f}}\left(\frac{\bar{\theta}}{\underline{\theta}^2}\right) + \delta$$

for arbitrarily small $\delta > 0$, may be derived using a considerably lengthier calculation.

It is straightforward to check whether the distribution of talents in the economy satisfies condition (15), hence whether the sufficient conditions given by Proposition 3 are satisfied. In the example below we show that *(15) is always satisfied under the uniform distribution.*

Example 3.1. Assume that the distribution of types is uniform, $\theta \sim U[\underline{\theta}, \bar{\theta}]$, then the term in curly brackets in the left hand side of (15) becomes

$$\frac{1}{2} - \frac{1}{2}\left(\frac{\underline{\theta}}{\theta}\right)^2,$$

which is clearly increasing in θ. An allocation solving the Mirrlees' Utilitarian program, $\mathcal{P}$, is therefore implementable.

Moreover, in this case the marginal tax rate is increasing, i.e., it is optimal to have a 'progressive' tax schedule.

The term progressive here refers to increasing *marginal* tax rates. Because marginal tax rates are non-positive for all agents total subsidies must be higher for more productive agents. Marginal subsidies do however decrease with earnings.

Proposition 3 leaves open the possibility that utilitarian solutions to Mirrlees' (1971) programs are always implementable, i.e., that a taxation principle always applies. It turns out that a necessary condition for implementation is for the firm's problem to be concave in $\dot{v}$ (see Lemma 7). Differentiating twice the planner's problem with respect to $\dot{v}$ we obtain $\phi''\left(-\dot{v}^*(\theta)\right) - \theta^2\chi''(v(\theta) - \theta\dot{v}(\theta))$. Using (13), we have

$$\chi''(u^*(\theta)) = \kappa''(u^*(\theta)) - \frac{\lambda}{\dot{u}^*(\theta)}\frac{d}{d\theta}\left[\frac{1}{f(\theta)\theta^2}\int_{\underline{\theta}}^{\theta} f(\tilde{\theta})\tilde{\theta}d\tilde{\theta}\right].$$

Hence, a necessary condition for the firm's problem to be concave in $\dot{v}$ is

$$\phi''\left(-\dot{v}^*(\theta)\right) - \theta^2\kappa''(u^*(\theta)) + \frac{\lambda\theta^2}{\dot{u}^*(\theta)}\frac{d}{d\theta}\left[\frac{1}{f(\theta)\theta^2}\int_{\underline{\theta}}^{\theta} f(\tilde{\theta})\tilde{\theta}d\tilde{\theta}\right] \leq 0. \tag{17}$$

This latter equation, (17), imposes a restriction on χ that depends on the local behavior of f, namely

$$\chi'(u(\theta)) \geq \frac{f'(\theta)}{2f(\theta)^2}\int_{\underline{\theta}}^{\theta} \chi'(u(a))f(a)da.$$

This condition need not be always satisfied. Next, we can offer an example of an environment for which it is not possible to implement the planner's utilitarian solution using a tax system.

Example 3.2. Let preferences be of the form

$$U(c, n, \theta) = \nu(c) - \theta n.$$

Then the Mirrlees's (1971) program becomes

$$\max \int_{\underline{\theta}}^{\bar{\theta}} \{h(\theta) - \kappa(u(\theta))\} f(\theta)d\theta$$

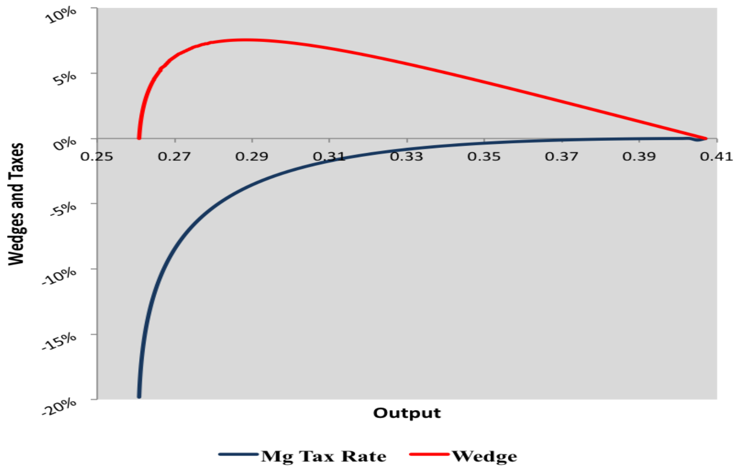

Figure 1: The figure displays the marginal tax rate, $\mathcal{T}'(\cdot)$, and the labor wedge, $\tau(\cdot)$, as a function of output, n.

subject to (2), (3), (4), and (5).

Let $\mathbb{E}[\theta] = \int_{\underline{\theta}}^{\overline{\theta}} \theta f(\theta) d\theta$. A little bit of algebra allows us to write the program above simply as

$$\max \int_{\underline{\theta}}^{\overline{\theta}} \left\{ u(\theta) \left[\frac{1}{\theta} + \frac{1}{\theta^2 f(\theta) \mathbb{E}[\theta]} \int_{\underline{\theta}}^{\theta} f(a) a da - \frac{F(\theta)}{\theta^2 f(\theta)} \right] - \kappa(u(\theta)) \right\} f(\theta) d\theta,$$

which solution is given by the first order condition,

$$\kappa'(u(\theta)) = \frac{1}{\theta} + \frac{1}{\theta^2 f(\theta)} \left[\frac{1}{\mathbb{E}[\theta]} \int_{\underline{\theta}}^{\theta} f(a) a da - F(\theta) \right], \tag{18}$$

and the monotonicity restriction that the term in the right hand side of (18) be increasing in θ.

As for the implementability constraint, $\chi'' \geq 0$, in this example it becomes

$$\frac{1}{\theta} + \frac{1}{\theta^2 f(\theta)} \left[\left(\frac{1}{\mathbb{E}[\theta]} - \lambda \right) \int_{\underline{\theta}}^{\theta} f(a) a da - F(\theta) \right] \tag{19}$$

increasing in θ.

Although similar, the monotonicity and the implementability conditions differ. It is then possible to choose a function $F(\cdot)$ such that the left hand side of (18) is increasing in θ, for all θ, but (19) is decreasing in an interval $[\theta_a, \theta_b] \in \Theta$ thus violating convexity of $\chi(\cdot)$.

3.3 Bunching

Up to this point we have dealt with the case in which the solution for the Mirrlees' (1971) program is such that $h(\cdot)$ is a strictly decreasing function of θ.

It is possible to generalize our results by assuming that the monotonicity restriction binds in the Mirrlees (1971) program, i.e. that there is bunching. Bunching introduces new challenges and technical issues for the application of our method. In particular, note that we rely on $\dot{u}^* < 0$ in (14) to assess implementability in our setting, which clearly cannot be true for all θ once we have intervals of bunching in the solution to the planner's problem. Moreover, it was this assumption that allowed us to focus on a relaxed program for the firm.

To deal with this difficulty, we consider a sequence of problems in which, for every n, bunching is allowed neither in the planner's program $[\mathcal{P}_n]$ nor in the firm's program $[\mathcal{P}_n^F]$. Formally, we consider a sequence of problems in which $-\dot{h}$ is restricted to a compact set $[n^{-1}, n]$. Under these restrictions, a suitable generalization of the method developed in the text applies. If we assume that (15) holds and, for each n, let $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ denote the solution to the planner's constrained problem $\mathcal{P}_n$, then we can construct a convex mapping $\chi_n(\cdot)$ that implements the solution $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ in the associated firm's problem $\mathcal{P}_n^F$. As n increases, the requirement that $-\dot{h}$ belongs to $[n^{-1}, n]$.becomes less stringent and consequently the sequence of solutions $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ converges [a.e.] to $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$, with the allocations converging [a.e.] to $(h^*(\theta), u^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$.

We show that the sequence of functions $\chi_n(\cdot)$ is equicontinuous and hence contains a subsequence that converges uniformly to a Lipshitz function, $\chi(\cdot)$. Finally, the convergence of $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ to $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ the convergence of $\chi_n(\cdot)$ to $\chi(\cdot)$ are used to assert that $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ solves the firm's program under $\chi(\cdot)$. Indeed, were it note the case $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ would not solve the firm's problem $\mathcal{P}_n^F$ for n large enough.

The whole argument is rather technical and we relegate it to appendix D. The main point is that our method works in this case as well.

4 Beyond Utilitarianism

In this section we generalize the objective of the planner. We focus on allocations that are associated with more redistributive metrics, in the sense of placing larger weights in more disadvantaged workers – higher θ.

We prove two results. First, we show that the more skewed the welfare weights are towards the less efficient workers, the easier it is for the associated sufficient condition for implementation to be satisfied. Second, as the planner's welfare metric approaches the Rawlsian limit the marginal tax rate faced by any type θ worker gets arbitrarily close to zero .

The allocations we investigate in this section solve a problem which differs from $\mathcal{P}$ by the fact that the planner attaches weights, $\omega(\theta)$, with $\omega(\theta) > 0$ for each type θ worker. That is, we let

$$\int_{\underline{\theta}}^{\bar{\theta}} \omega(\theta)v(\theta)f(\theta)d\theta \geq w \tag{20}$$

substitute for (5) in program $\mathcal{P}$. We consider a sequence of Pareto problems that approach a Rawlsian limit in which only the utility of the worse of worker matters. Toward this end we construct a sequence of Programs, $\tilde{\mathcal{P}}_n$ with associated weights $\omega_n(\theta)$ of the form $\omega_n(\theta) = A_n\theta^{n-1}$, where A_n is chosen in such a way as to guarantee that, for all n, $\int_{\underline{\theta}}^{\bar{\theta}} \omega_n(\theta)f(\theta)d\theta = 1$. Note that the relative weights $\omega_n(\theta)/\omega_n(\bar{\theta})$ converge to 0 as n goes to infinity.

An analogous derivation to the one used in Section 3.2, allows us to derive for

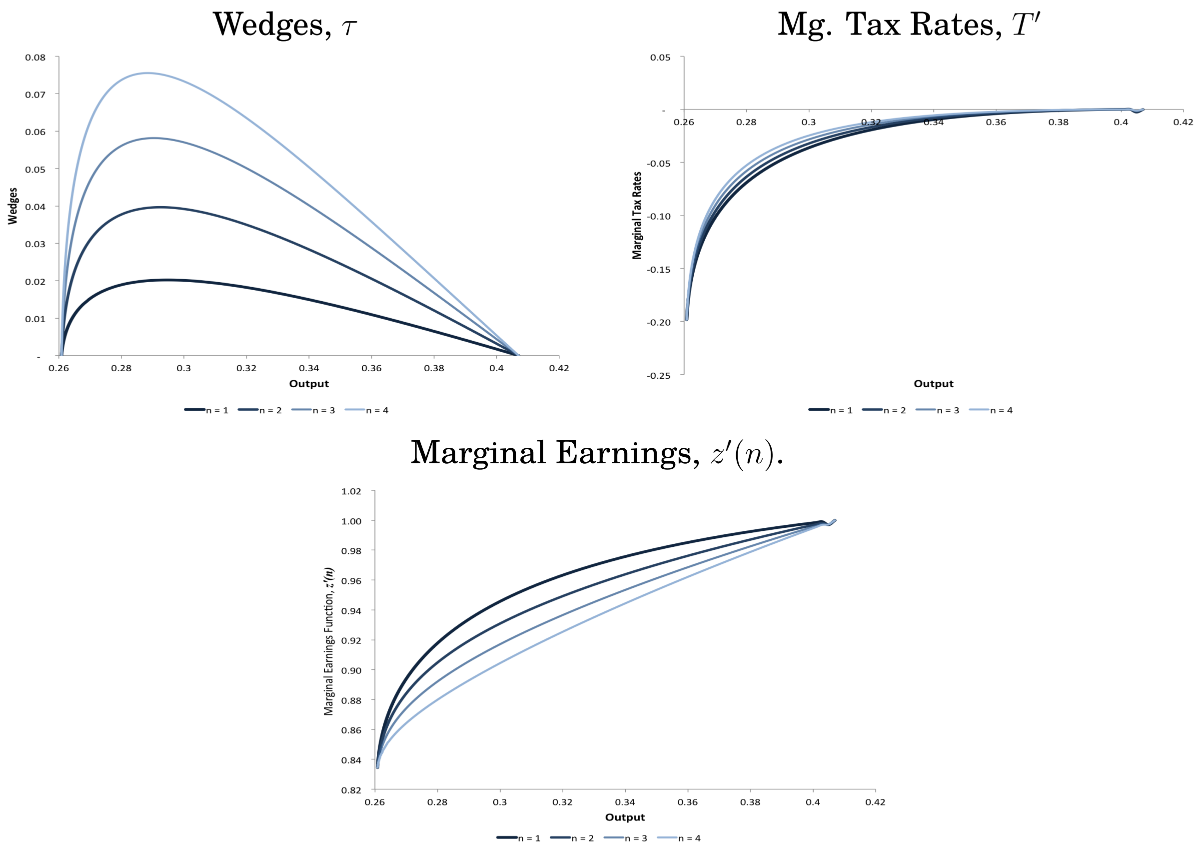

Figure 2: **Non-utilitarian Metrics:** The figures display wedges, τ, marginal taxe rates, $\mathcal{T}'$, and marginal earnings, z', for different distributions of Pareto weights in the Mirrlees' (1971) program. For each n, the Pareto weights are $\omega(n) = A_n \theta^{n-1}$, $\int \omega_n(\theta) f(\theta) d\theta = 1, \forall n$.

each n the sufficient condition (15) for implementation,

$$\frac{d}{d\theta}\left\{\frac{1}{\theta^2 f(\theta)}\int_{\underline{\theta}}^{\theta}\omega_n(a) a f(a) da\right\} \geq 0\,,\,\forall\theta. \tag{21}$$

For each n, let $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ be the solution $\tilde{\mathcal{P}}_n$ and $(h_n^*(\theta), u_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$, the associated optimal allocation. We assume that for all n, $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ satisfies Assumption 1. Additionally, we assume that the sequence of Lagrange multipliers, $(\lambda_n)_n$, associated with the isoperimetric constraints, (20), for each $\tilde{\mathcal{P}}_n$ is uniformly bounded by some constant $\lambda > 0$. This conditions always holds, for instance, when labor supply belongs to a bounded set.

Next we show that when the Pareto weights are sufficiently biased towards the least talented workers, the planner's preferred allocations is always implementable by a tax system.

Proposition 4. *For all n, let the allocation that solves the Planner's program, $\tilde{\mathcal{P}}_n$, satisfy Assumption 1. Then, there is $n^* \in \mathbb{N}$ such that for all $n > n^*$ the allocation is implementable.*

Proof. For any n the sufficient condition for implementation is, in this case,

$$\frac{d}{d\theta}\left\{\frac{1}{\theta^2 f(\theta)}\int_{\underline{\theta}}^{\theta}\omega_n(a) a f(a) da\right\} = A_n \frac{\theta^{n+2} f(\theta)^2 - [2\theta f(\theta) + \theta^2 f'(\theta)]\int_{\underline{\theta}}^{\theta} a^n f(a) da}{\theta^4 f(\theta)^2}.$$

Let $\epsilon^f(\theta) := \theta f'(\theta)/f(\theta)$, $\bar{\epsilon}^f := \max_\theta \epsilon^f(\theta)$ and $\bar{f} = \max_\theta f(\theta)$. Then, it is not hard to see that (21) holds whenever

$$f(\theta)\theta^{n+2} \geq \theta\left[2+\bar{\epsilon}^f\right]\frac{\bar{f}}{n+1}\left[\theta^{n+1}-\underline{\theta}^{n+1}\right] \ \forall\theta.$$

It follows that the last inequality always holds for

$$n \geq \left(2+\bar{\epsilon}^f\right)\frac{\bar{f}}{\underline{f}} - 1.$$

□

According to Proposition 4, $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ is implementable for large enough n. For each n let $\chi_n(\cdot)$ be a function which defines the firms' programs solved by $(v_n^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$.[12] Proposition below shows that as n increases $\chi_n'(u_n^*(\theta))$ gets arbitrarily close to $\kappa'(u_n^*(\theta))$ for all $\theta < \bar{\theta}$. That is, as n increases and $\tilde{\mathcal{P}}_n$ approaches the Rawlsian limit, marginal tax rates converge to zero for every type $\theta < \bar{\theta}$.

Proposition 5. *Assume that for all n the allocation which solves the planner's program, $\tilde{\mathcal{P}}_n$, satisfies Assumption 1. Then, for every $\varepsilon > 0$ there exists N such that $n > N$ implies $|\chi_n'(u_n^*(\theta)) - \kappa'(u_n^*(\theta))| < \varepsilon$ for every $\theta < \bar{\theta} - \varepsilon$.*

Proof. For all θ,

$$\chi_n'(u_n(\theta)) - \kappa'(u_n(\theta)) = \frac{\lambda_n}{\theta^2 f(\theta)}\int_{\underline{\theta}}^{\theta}\omega_n(a)af(a)da.$$

Because $\int_{\underline{\theta}}^{\bar{\theta}}\omega_n(\theta)f(\theta)d\theta = 1$, we must have

$$A_n\underline{f}\int_{\underline{\theta}}^{\bar{\theta}}\theta^{n-1}d\theta \leq 1 \Rightarrow A_n \leq \frac{1}{\underline{f}}\frac{n}{\bar{\theta}^n-\underline{\theta}^n}$$

For n sufficiently large we may further simplify to $A_n \leq 2n\bar{\theta}^{-n}/\underline{f}$. Then, since $\lambda_n \leq \bar{\lambda}$ we have

$$\begin{aligned}\chi_n'(u_n(\theta)) - \kappa'(u_n(\theta)) &\leq \frac{\bar{\lambda}}{\theta^2 f(\theta)}\int_{\underline{\theta}}^{\theta}A_n a^n f(a)da \\ &\leq \frac{\bar{\lambda}}{\theta^2 f(\theta)}\frac{2n}{\bar{\theta}^n\underline{f}}\int_{\underline{\theta}}^{\theta}a^n f(a)da = 2\bar{\theta}\frac{\bar{\lambda}}{\underline{f}}\frac{n}{n+1}\frac{\bar{f}}{\underline{f}}\left(\frac{\theta}{\bar{\theta}}\right)^{n-1}.\end{aligned}$$

Hence, $\chi_n'(u_n(\theta)) - \kappa'(u_n(\theta)) \to 0$ as $n \to \infty$. □

The intuition behind Propositions 4 and 5 is straightforward. As the planner's preferences approach the Rawlsian metric, it places less and less weight on types $\theta < \bar{\theta}$. In the limit, its program becomes maximizing the economy's resources subject to giving a minimum utility to the least efficient worker. But this problem is equivalent to the firm's problem. Therefore incentives are aligned and marginal tax rates should be equal to zero. Figure 2, displays the effect on wedges and taxes of increasing n.

[12] As we have seen above, the multiplicity stems from the level of subsidy for the least talented worker.

5 Alternative Mechanisms

Our restriction to tax schedules or, more generally, tax systems borrows from the results in Hammond (1979, 1987) that show the equivalence between mechanisms and tax systems in competitive markets – see Guesnerie (1998). This equivalence fails, however, to hold in some environments when, for example, the anonymity assumption that underlies most implementations via tax schedules is relaxed, e.g., Piketty (1993). In this section we ask whether we can implement more by assuming that the monopsonistic power is caused by the presence of a single firm in the market — 5.1.

Next, back to our baseline rationale for market power, we show how crossing information of firms and workers and blurring the distinction between workers and firms by conditioning a worker's allocations on profits of the firm he works for may increase the set of implementable allocations — 5.2.

5.1 Monopsony

Thus far we have exploited the similarity between each firm's program and that of a monopsonistic firm. That is, the expected profit maximization problem of each firm is formally equivalent to the profit maximization problem of a single monopsonistic firm. The question we address here is whether one can do more in the case of a true monopsony than in the economy we have been studying.

We first note that each agent's earnings are assumed to be observed by the planner. Otherwise, the planner would not be able to impose any tax schedule. Since there is only one firm in the economy, observing the distribution of earnings for the whole economy is the same as observing the distribution of payments made by the firm. Because earnings map one to one into consumption and because incentive compatibility requires monotonicity, the planner is able to recover the whole profile, $(u(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ from the economy's distribution of earnings.

Consider then the following mechanism. For a worker who announces to be of type $\underline{\theta}$ the planner offers an unemployment benefit that yields a utility $v^*\left(\bar{\theta}\right)$ if he does not work. To the firm, the planner suggests that it implements the consumption profile, $c^*(\theta) = \kappa(u^*(\theta))$, $\forall\theta$, $(u^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$, the consumption profile corresponding to the Mirrlees' (1971) problem's solution. If it the firm does as suggested, the planner requires a transfer from (or makes a transfer to) the firm that leads to zero profits. If it does not, the planner demands a payment equal to ∞. Extreme punishment guarantees obedience.

We prove next that in the unique best-response of the firm it chooses an allocation which is [a.e.] equal to the one induced by the planner's problem.

Proposition 6. *A mechanism that uses information on the entire distribution of earnings implements the Mirrlees' (1971) allocation, if the planner is allowed to use strong punishment on the firm.*

Proof. In any solution to the firm's problem $(c(\theta), h(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ the least efficient worker, $\bar{\theta}$ must obtain zero rent, as we have seen. Hence, $\nu(c\left(\bar{\theta}\right)) - \bar{\theta}h(\bar{\theta}) = v^*\left(\bar{\theta}\right)$. Incentive-compatibility implies monotonicity which implies that the firm should give the lowest consumption to $\bar{\theta}$. If the firm implements the consumption profile suggested by the government, it must be the case that $c(\bar{\theta}) = c^{*}(\bar{\theta})$, which further implies $h(\bar{\theta}) = h^{*}(\bar{\theta})$. Moreover, also from incentive compatibility $\dot{v}(\theta) = -h(\theta)$ [a.e.], leading to

$$\int_{\theta}^{\bar{\theta}} d\nu(c(\tilde{\theta})) - \int_{\theta}^{\bar{\theta}} dh(\tilde{\theta}) = 0,$$

where the Riemann-Stieltjes integrals above are well defined because both $\nu \circ c$ and h are monotonic.

Therefore, for almost all θ, we have

$$h(\theta) = h(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} dh(\tilde{\theta}),$$

which is equal to $h^{*}(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} d\nu(c(\tilde{\theta}))$.

Since the firm chooses the same consumption distribution as the government, the latter term equals

$$h^{*}(\bar{\theta}) - \int_{\theta}^{\bar{\theta}} d\nu(c^{*}(\tilde{\theta})) = h^{*}(\theta)$$

for almost all θ. □

The message we get from this section is that, if the planner can use information on the entire consumption profile and a policy which imposes strong punishment for a failure to deliver a desired profile, it may in practice implement any incentive-feasible allocation it desires. We reached this conclusion by treating the firm as if it were an agent endowed with a well defined objective. This deserves some qualifications, however, since the firm's objective must be ultimately derived from that of its owners. In particular, we said nothing about the firm's ownership structure in this setting and whether profit maximization should still be derived as its objective.

If we assume a diluted ownership as we have been assuming for the 'small' firms we have been considering so far, the fact that the firm has monopsonisitc power means that we can no longer guarantee that agents agree on profit maximization as its goal. Indeed, the set of rules under which decision is made — majority voting, qualified voting, etc. — becomes key in this regard. Another possibility is that a firm is owned by a single agent. A non-anonymous mechanism that singles out this agent by requiring him to deliver the desired utility profile is what we are prescribing, in this case. The non-anonymous nature of the mechanism stems from the use of one worker's identity as the owner of the firm to, through extreme punishment, induce the desired allocation.

However consistent with the underlying informational structure of the problem, the suggested implementation relies too heavily on the fact that whole distribution of earnings is generated by a single firm.[13]

Having observed that the there are cases in which the taxation principle fails, we ask what kind of additional information helps the firm. Section 5.1.1 shows that even if the planner knows the total amount spent by the firm on wages it will not to be able to implement its desired allocation.

[13] Such implementation is not, for example, possible if we consider the random matching interpretation of our setting.

5.1.1 Information on Aggregate Payments

If we maintain the assumption of a true monopsony can we enlarge the set of implementable allocations from that of small firms randomly matched with workers but still using tax systems? I.e., let us enrich the set of instruments as compared to that offered by small firms by assuming that the firm faces an additional tax on its total payments,

$$\int_{\underline{\theta}}^{\bar{\theta}} \chi\left[v(\theta) - \theta\dot{v}(\theta)\right] f(\theta)d\theta.$$

Clearly, we can assume that this tax is 0 if the payments are equal to $C := \int_{\underline{\theta}}^{\bar{\theta}} \chi\left[v^*(\theta) - \theta\dot{v}^*(\theta)\right] f(\theta)d\theta$ and ∞ otherwise. In this case, we can use Lemma (1) to write the firm's problem as an isoperimetric problem.[14] There is $\beta \in \mathbb{R}$ for which the firm maximizes

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(-\dot{v}(\theta)) - \chi(v(\theta) - \theta\dot{v}(\theta))\right] f(\theta)d\theta + \beta\left[\int_{\underline{\theta}}^{\bar{\theta}} \chi\left[v(\theta) - \theta\dot{v}(\theta)\right] f(\theta)d\theta - C\right].$$

Letting $\hat{\chi} := \chi(1-\beta)$ we write the firms's problem as

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(-\dot{v}(\theta)) - \hat{\chi}(v(\theta) - \theta\dot{v}(\theta))\right] f(\theta)d\theta - \lambda C.$$

This problem above is identical to program $\mathcal{P}^F$ when we replace χ by $\hat{\chi}$. Proposition 7, below, ensues.

Proposition 7. *If the planner's desired allocation $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ is not implementable when the government does not use information on the aggregate payments, then it remains non-implementable when the government can add a tax on the firm's total aggregate payments.*

5.2 Profit-dependent Taxes on Earnings

Back to our random matching assumption, consider the specific case in which each firm is matched with a single agent. If the planner can cross the information about each worker's earnings and the profits earned by the firm he works for, the planner can in practice recover all information about the output that the worker is producing.

Hence, a mechanism under which the allocation assigned to the worker depends not only on his earnings but also on the profits of the firm he works for may, for all purposes, condition on the output he produces. This type of mechanism would allow the implementation of any Mirrlees' (1971) allocation.

Of course, the dependence of earnings taxation on firms' profits blurs the traditional separation between a worker and a firm. Besides, perfect identification of output is only possible in the case of one worker for each firm. Still, at least from a theoretical perspective, the use of such mechanisms are likely to increase the set of implementable allocations.

[14] We apply Theorem 17.9 in Clarke (2013)

6 Conclusion

In this paper we take a first step towards bringing the analysis of optimal tax policy under imperfect competition to the same footing as that under competitive markets. We address optimal tax policy in what is effectively a monopsonistic labor market using a mechanism design approach.

The main driving force of our results is the relative value placed by the planner on highly productive agents. Participation constraints force a firm with market power to provide enough utility for the lowest type it hires. For any type above, the only thing that constrains the firm is incentive compatibility. No direct value of increasing one's welfare is warranted.

If market power is not so extreme, the firm will have to worry about losing workers to competition and will most probably be bound by these constraints. Clear cut results as the one found for this polar case are unlikely to arise. Yet, we conjecture that the forces that we have uncovered here will still be important there.

References

David P. Baron and Roger B. Myerson. Regulating a monopolist with unknown costs. *Econometrica*, 50:911 – 930, 1982.

Pierre Cahuc and Guy Laroque. Optimal taxation and monopsonistic labor market: Does monopsony justify the minimum wage? *Journal of Public Economic Theory*, 16(2):259–273, 2014.

Francis Clarke. *Functional Analysis, Calculus of Varaiations and Optimal Control*. Springer, 2013.

Roger Guesnerie. *A Contribution to the Pure Theory of Taxation*. Cambridge University Press, 1998.

Peter J. Hammond. Straightforward individual incentive compatibility in large economies. *Review of Economic Studies*, 46(0):263–282, 1979.

Peter J Hammond. Markets as constraints: Multilateral incentive compatibility in continuum economies. *Review of Economic Studies*, 54(3):399–412, 1987.

Alan Manning. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press, 2005.

James A. Mirrlees. An exploration in the theory of optimal income taxation. *Review of Economic Studies*, 38:175–208, 1971.

Thomas Piketty. Implementation of first-best allocations via generalized tax schedules. *Journal of Economic Theory*, 61:23–41, 1993.

Robert B. Reich. *Saving Capitalism: For the Many, Not the Few*. Alfred A. Knopf, 1st edition, 2015.

Jesus Seade. On the shape of optimal tax schedules. *Journal of Public Economics*, 7:203–236, 1977.

Stephanie Stantcheva. Optimal income taxation with adverse selection in the labour market. *Review of Economic Studies*, 81:1296 – 1329, 2014.

A Mathematical Appendix

B Proofs

Proof of Lemma 3. Let

$$\Lambda\left(\theta, v(\theta), \dot{v}(\theta)\right) := \left[\phi(-\dot{v}(\theta)) - \chi(v(\theta) - \theta\dot{v}(\theta))\right] f(\theta),$$

and define for any absolutely continuous $v \in AC[\underline{\theta}, \overline{\theta}]$ (with derivitve $\dot{v}$),

$$\Pi\left(v\right) = \int_{\underline{\theta}}^{\bar{\theta}} \Lambda\left(\theta, v(\theta), \dot{v}(\theta)\right) d\theta.$$

We restrict our attention to functions v such that $v(\bar{\theta}) \in \left[v^*(\bar{\theta}), u_2\right]$ – recall that $\mathbb{U} := [u_1, u_2]$ and $u_2 > u\left(c^*(\underline{\theta})\right)$. We would like to apply Theorem 16.2 and Exercise 16.8 from Clarke (2013) to guarantee that the problem admits a solution in the class of the absolutely continuous functions, $AC[\underline{\theta}, \overline{\theta}]$, such that $v(\bar{\theta}) \in \left[v^*(\bar{\theta}), u_2\right]$.

Take H such that $u_2 - \underline{\theta}H = v^*\left(\bar{\theta}\right)$. We define $\phi(h) = -\infty$ for all $h \notin [0, H]$ and $\chi(u) = \infty$ for all $u \notin \mathbb{U}$. Under these restrictions, if we guarantee that (ϕ, χ) is concave when restricted to the domain $[0, H] \times \mathbb{U}$, it will also be concave when its domain is $\mathbb{R}^2$.

Notice that the objective functions is trivially coercitive of degree 2 with respect to $\dot{v}$ (as it is uniformly bounded above).

It is also concave in $\dot{v}$ by assumption. The only assumption from Theorem 16.2 and Exercise 16.8 which does not apply is the one that requests that $\Lambda\left(\theta, v(\theta), \dot{v}(\theta)\right)$ is continuous. Hence we show how to modify the proof of these Theorems to deal with this case. For that let $\bar{\Pi}$ be the supremum of the profit function over $AC[\underline{\theta}, \overline{\theta}]$ and take a sequence $v_n \in AC[\underline{\theta}, \overline{\theta}]$ such that

$$\bar{\Pi} = \lim \Pi\left(v_n\right).$$

We may thus assume that $\dot{v}_n \in [0, H]$ almost everywhere (otherwise $\phi(\dot{v}_n\left(\theta\right)) = -\infty$ in a set of positive measure and thus $\Pi\left(v_n\right) = -\infty$). Therefore since $\dot{v}_n \in [0, H]$ a.e. we may apply Alouglu Theorem (Theorem 3.14 in Clarke (2013)) to find a subsequence (again without relabelling) and a function $\left(r\left(\theta\right)\right)_{\theta \in [\underline{\theta}, \bar{\theta}]}$ in $L^2[\underline{\theta}, \overline{\theta}]$ for which $(\dot{v}_n)$ converges weakly in $L^2[\underline{\theta}, \overline{\theta}]$ to r. Also taking a sunsequence if necessary we may assume that the sequence $v_n(\bar{\theta})$ is convergent and write $\hat{v}(\bar{\theta})$ for its limit. Clearly we have $\hat{v}(\bar{\theta}) \in \left[v^*(\bar{\theta}), u(c^*(\underline{\theta}))\right]$.

As in the proof of Theorem 16.2 in Clarke (2013), notice that for all $\theta \in [\underline{\theta}, \overline{\theta}]$ the indicator function of each interval $[\theta, \bar{\theta}]$, $\mathbb{I}_{[\theta, \bar{\theta}]}$, lies in the dual of $L^2[\underline{\theta}, \overline{\theta}]$ and hence since $(\dot{v}_n)$ converges weakly to r we have

$$\int \mathbb{I}_{[\theta, \bar{\theta}]} r\left(\tilde{\theta}\right) d\tilde{\theta} = \lim \int \mathbb{I}_{[\theta, \bar{\theta}]} \dot{v}_n\left(\tilde{\theta}\right) d\tilde{\theta}.$$

Also following the proof of Theorem 16.2 in Clarke (2013) we define $\left(\hat{v}\left(\theta\right)\right)_{\theta \in [\underline{\theta}, \bar{\theta}]}$ by

$$\hat{v}(\theta) := \hat{v}(\bar{\theta}) - \int \mathbb{I}_{[\theta, \bar{\theta}]} r\left(\tilde{\theta}\right) d\tilde{\theta}.$$

It is straightforward to check that $\hat{v}(\theta)$ is continuously differentiable and that $\dot{\hat{v}}_n(\theta) = r(\theta)$ a.e.

Finally notice that Λ is upper semicontinuous with respect to $(v, \dot{v})$. Hence Hypothesis 6.37 in Clarke (2013) applies and Theorem 6.38 in Clarke (2013) imply that

$$\bar{\Pi} = \lim \int_{\underline{\theta}}^{\theta} \Lambda\left(\theta, v_n(\theta), \dot{v}_n(\theta)\right) d\theta \leq \int_{\underline{\theta}}^{\theta} \Lambda\left(\theta, \hat{v}(\theta), \dot{\hat{v}}_n(\theta)\right) d\theta,$$

which shows that $\hat{v}$ is a solution. □

Proof of Proposition 2. Consider the optimal allocation (v^*) and a policy of excluding all types $\theta \in [\bar{\theta}-\varepsilon, \bar{\theta}]$. Define $\theta_\varepsilon \in [\bar{\theta}-\varepsilon, \bar{\theta}]$ through

$$\int_{\bar{\theta}-\varepsilon}^{\bar{\theta}} \left(\phi(h^*(u)) - \chi(u^*(a))\right) f(a)da = \varepsilon f(\theta_\varepsilon)\left(\phi(h^*(\theta_\varepsilon)) - \chi(u^*(\theta_\varepsilon))\right).$$

The policy of exluding all types $\theta \in [\bar{\theta}-\varepsilon, \bar{\theta}]$ leads the firm to forgo a total of

$$f(\theta_\varepsilon)\left(\phi(h^*(\theta_\varepsilon)) - \chi(u^*(\theta_\varepsilon))\right)$$

in profits.

Next, notice that the utility received by the $\bar{\theta}-\varepsilon$ in the planner's solution satisfies $v^*(\bar{\theta}-\varepsilon) \geq v^*(\bar{\theta}) + \varepsilon h^*(\bar{\theta})$. Therefore the change above allows the firm to increase h uniformly for all types $\theta \in [\underline{\theta}, \bar{\theta}-\varepsilon]$ by at least $\varepsilon h^*(\bar{\theta})/(\bar{\theta}-\varepsilon)$. This increases profits by no less than

$$\frac{\varepsilon h^*(\bar{\theta})}{\bar{\theta}-\varepsilon} \int_{\underline{\theta}}^{\bar{\theta}-\varepsilon} \phi'\left(h^*(a) + \frac{\varepsilon h^*(\bar{\theta})}{\bar{\theta}-\varepsilon}\right) f(a)da,$$

We have used the concavity of ϕ to obtain the lower bound above. Because this must be true for all $\varepsilon > 0$, we obtain the following necessary condition for $\chi(u^*(\theta))$,

$$\chi(u^*(\bar{\theta})) \leq \phi(h^*(\bar{\theta})) - \frac{h^*(\bar{\theta})}{\bar{\theta}} \frac{1}{f(\bar{\theta})} \int_{\underline{\theta}}^{\bar{\theta}} \phi'(h^*(\theta)) f(\theta) d\theta.$$

For the rest of the proof, just add $-\kappa(u^*(\bar{\theta}))$ in both sides (B) and use, from Proposition 1, the fact that $\kappa'(u^*(\theta)) \geq \chi'(u^*(\theta))$ for all θ. □

Proof of Proposition 3. We will show that if the subsidy is

$$\chi(u^*(\bar{\theta})) - \phi\left(h^*(\bar{\theta})\right) = -\left[\phi\left(h^*(\bar{\theta})\right) - \phi\left(\frac{h^*(\bar{\theta})}{2}\right)\right] - \phi'\left(\frac{h^*(\bar{\theta})}{2}\right)\frac{h^*(\bar{\theta})}{\underline{f}}\left(\frac{\bar{\theta}}{\underline{\theta}^2}\right), \tag{22}$$

we can choose marginal tax schedules for a tax system which implements the planner's allocation.

Let $\|\chi'\| := \max_{u\in\mathbb{U}} \chi'$, and take a small

$$\varrho \in \left(0, \min\left\{\frac{h^*(\bar{\theta})\bar{\theta}}{4}; \|\chi'\|^{-1}\left[\phi\left(\frac{3h^*(\bar{\theta})}{4}\right) - \phi\left(\frac{h^*(\bar{\theta})}{2}\right)\right]\right\}\right). \tag{23}$$

Let the set of possible utilities that each firm is capable of delivering be $\mathbb{U} := [u^*(\bar{\theta})-\varrho, u^*(\underline{\theta})+\varrho]$. Define $\chi(\cdot)$ on $[u^*(\bar{\theta}), u^*(\underline{\theta})]$ using (13), (22), and the boundary condition, $\phi'(-\dot{v}^*(\underline{\theta})) = \underline{\theta}\chi'(v^*(\underline{\theta})-\underline{\theta}\dot{v}^*(\underline{\theta}))$. Next, extend $\chi(\cdot)$ continuously to $[u^*(\bar{\theta})-\varrho, u^*(\underline{\theta})+\varrho]$ by making $\chi'(\cdot)$ constant in $[u^*(\bar{\theta})-\varrho, u^*(\bar{\theta})]$ and in $[u^*(\underline{\theta}), u^*(\underline{\theta})+\varrho]$.

If we guarantee that the firm's problem is concave in $(v, \dot{v})$ then the Euler and transversality conditions guarantee that if the firm chooses to exclude no type then $(v^*(\theta))_{\theta\in[\underline{\varrho},\bar{\theta}]}$ solves the firm's relaxed problem when monotonicity constraints are ignored – see Theorem 18.8 in Clarke (2013). Therefore it suffices to show that it is not optimal for the firm to exclude any type.

Consider any mechanism proposed by a firm in which it hires all types $\theta \leq \hat{\theta}$ for some $\hat{\theta} < \bar{\theta}$, but no types $\theta > \hat{\theta}$. We show that the firm can strictly increase its profits by hiring a larger set of workers (a set $[\underline{\theta}, \hat{\theta}+\varepsilon]$ for some $\varepsilon > 0$).

Since a type $\hat{\theta}$ worker consumes at least $u^*(\bar{\theta})-\varrho$ and since he is indifferent between being hired and not being hired he obtains utility $v^*(\bar{\theta})$. Therefore he must incur a disutility of effort weakly greater than $\breve{h}(\hat{\theta})$ where, for any θ', $\breve{h}(\theta')$ is implicitly defined by $u^*(\bar{\theta})-\varrho-\theta'\breve{h}(\theta') = v(\bar{\theta})$.

Consider now a deviation in which the firm also hires the types $\theta \in [\hat{\theta}+\varepsilon, \hat{\theta})$ by offering them $\left(u^*(\bar{\theta})-\varrho, \breve{h}(\hat{\theta}+\varepsilon)\right)$. A lower bound on the change in expected profits from adding these types, without considering the need to change the more efficient types' allocations, is

$$\varepsilon\underline{f}\left[\phi(\breve{h}(\hat{\theta}+\varepsilon))-\chi(u^*(\bar{\theta})-\varrho))\right] \geq \varepsilon\underline{f}\left[\phi(\breve{h}(\hat{\theta}+\varepsilon))-\chi(u^*(\bar{\theta}))-\|\chi'\|\,\varrho\right]. \tag{24}$$

Notice that

$$\breve{h}(\hat{\theta}+\varepsilon) \geq \breve{h}(\bar{\theta}) = \frac{u^*(\bar{\theta})-v(\bar{\theta})}{\bar{\theta}} - \frac{\varrho}{\bar{\theta}} > \frac{3h^*(\bar{\theta})}{4},$$

by (23). Hence (41) is strictly more than

$$\varepsilon\underline{f}\left[\phi\left(\frac{3h^*(\bar{\theta})}{4}\right)-\chi(u^*(\bar{\theta}))-\|\chi'\|\,\varrho\right]. \tag{25}$$

Also by (23) $\varrho < \left(\frac{1}{\|\chi'\|}\right)\left[\phi\left(\frac{3h^*(\bar{\theta})}{4}\right)-\phi\left(\frac{h^*(\bar{\theta})}{2}\right)\right]$ and hence (25) is more than

$$\varepsilon\underline{f}\left[\phi\left(\frac{h^*(\bar{\theta})}{2}\right)-\chi(u^*(\bar{\theta}))\right]. \tag{26}$$

Notice that the new allocation is monotonic but not incentive compatible, as type $\hat{\theta}$ would prefer to mimic all types $\theta \in [\hat{\theta}+\varepsilon, \hat{\theta})$. Indeed this choice would deliver

$$\begin{aligned} u^*(\bar{\theta})-\varrho-\hat{\theta}\breve{h}(\hat{\theta}+\varepsilon) &= u^*(\bar{\theta})-\left(\hat{\theta}+\varepsilon\right)\breve{h}(\hat{\theta}+\varepsilon)+\varepsilon\breve{h}(\hat{\theta}+\varepsilon) \\ &= v(\bar{\theta})+\varepsilon\breve{h}(\hat{\theta}+\varepsilon). \end{aligned}$$

Notice that to make the new allocation incentive compatible it suffices to decrease h for all $\theta \leq \hat{\theta}$ uniformly by $\frac{\varepsilon\breve{h}(\hat{\theta}+\varepsilon)}{\hat{\theta}}$. A strict upper bound on the loss due to this change is

$$\int_{\underline{\theta}}^{\bar{\theta}}\left\{\phi\left(h_o(\theta)\right)-\phi\left(h_o(\theta)-\frac{\varepsilon\breve{h}(\hat{\theta}+\varepsilon)}{\hat{\theta}}\right)\right\}f(\theta)d\theta < \phi'(\breve{h}(\bar{\theta}))\frac{\varepsilon\breve{h}(\hat{\theta}+\varepsilon)}{\hat{\theta}}, \tag{27}$$

where $h_o(\theta)$ is the candidate solution for the firm's program and where we used the fact that ϕ is concave to obtain the upper bound above. Since $\frac{\varepsilon\breve{h}(\hat{\theta}+\varepsilon)}{\hat{\theta}} < \frac{\varepsilon\breve{h}(\underline{\theta})}{\underline{\theta}}$ and using (23) we have

$$\breve{h}(\bar{\theta}) = \frac{u^*(\bar{\theta}) - v(\bar{\theta})}{\bar{\theta}} - \frac{\varrho}{\bar{\theta}} = h^*(\bar{\theta}) - \frac{\varrho}{\bar{\theta}} > \frac{3h^*(\bar{\theta})}{4} > \frac{h^*(\bar{\theta})}{2}.$$

Hence, the left hand side of (27) is less than $\phi'\left(\frac{h^*(\bar{\theta})}{2}\right)\frac{\varepsilon\breve{h}(\underline{\theta})}{\underline{\theta}}$. Since

$$\breve{h}(\underline{\theta}) = \left(\frac{u^*(\bar{\theta}) - v(\bar{\theta})}{\underline{\theta}} - \frac{\varrho}{\underline{\theta}}\right) \leq \left(\frac{\bar{\theta}}{\underline{\theta}}\right) h^*(\bar{\theta}),$$

the left hand side of (27) is less than

$$\phi'\left(\frac{h^*(\bar{\theta})}{2}\right)\frac{\varepsilon\bar{\theta}h^*(\bar{\theta})}{\underline{\theta}^2}. \tag{28}$$

Therefore a sufficient condition for no exclusion is that the lower bound for the gains (26) is strictly greater than the upper bound for the losses (28):

$$\varepsilon\underline{f}\left[\phi\left(\frac{h^*(\bar{\theta})}{2}\right) - \chi(u^*(\bar{\theta}))\right] \geq \phi'\left(\frac{h^*(\bar{\theta})}{2}\right)\frac{\varepsilon\bar{\theta}h^*(\bar{\theta})}{\underline{\theta}^2},$$

which holds by the construction of χ.

□

Proof of Lemma 2. Immediate from Lemmas 4 and 6. □

C Lemmata

In all that follows it will be convenient to define

$$\Lambda(\theta, v, \dot{v}) := [\phi'(-v''(v - \theta\dot{v})]\, f(\theta).$$

We start with the following lemma.

Lemma 4. *If $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ solves the firm's problem, $\mathcal{P}^F$, then the Euler equation (8) holds a.e.*

Proof. Since we are assuming that χ is C^1 the functional $\Lambda(\theta, v, \dot{v})$ is C^1 in a neighborhood of $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$.[15] Consider a particular Lipschitz function $(y(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ such that $y(\underline{\theta}) = y(\bar{\theta}) = 0$. Now, take a scalar λ and consider the functional

$$g(\lambda) := \int_{\underline{\theta}}^{\bar{\theta}} \Lambda(\theta, v^*(\theta) + \lambda y(\theta), \dot{v}^*(\theta) + \lambda\dot{y}(\theta))\, f(\theta)d\theta.$$

[15]For two AC functions $(v^1(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ and $(v^2(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ with uniformly bounded derivatives $v^{i\prime}(\theta)$ (for $i = 1, 2$) we consider the metric

$$d(v^1, v^2) = \sup_{\theta} \max\left\{\left|v^1(\theta) - v^2(\theta)\right|, \left|v^{1\prime}(\theta) - v^{2\prime}(\theta)\right|\right\}.$$

Notice that for any Lipschitz $(y(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ there exists $\varepsilon>0$ such that $|\lambda|<\varepsilon$ implies $\dot{v}^*(\theta)+\lambda y'(\theta)<0$ for all θ. Thus, the function $(v^*(\theta)+\lambda y(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ is monotonic. Hence a necessary condition for optimality is $g'(0)=0$. The argument from Theorem 15.2 in Clarke (2013) implies the desired result. □

Lemma 5. *The function χ and hence the functional Λ are smooth around a neighborhood of $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$.*

Proof. From (8) we have

$$\theta f(\theta)\frac{d}{d\theta}\{\chi'(v^*(\theta)-\theta\dot{v}^*(\theta))\}=-2\chi'(v^*(\theta)-\theta\dot{v}^*(\theta))f(\theta)+$$
$$\theta\chi'(v^*(\theta)-\theta\dot{v}^*(\theta))f''(\theta)+\phi''(-\dot{v}^*(\theta)))v^{**}(\theta)-\phi'(-\dot{v}^*(\theta))f''(\theta). \quad (29)$$

Since v^* is smooth and χ is C^1 by assumption, the right hand side of (40) is continuous. Hence, $\frac{d}{d\theta}\{\chi'(v^*(\theta)-\theta\dot{v}^*(\theta))\}$ is C^1, which implies that χ is C^2. This in turn, implies that the right hand side of (40) is C^1 which, by a similar argument, implies that χ is C^3. Repeating the same argument we conclude that $\chi\in C^k$ for all $k\in\mathbb{N}$. We conclude that $\Lambda(\theta,v,\dot{v})$ is smooth in a neighbourhood of $(v^*(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$. □

Lemma 6. *The transversality condition, $\phi'(-\dot{v}^*(\underline{\theta}))=\underline{\theta}\chi'(v^*(\underline{\theta})-\underline{\theta}\dot{v}^*(\underline{\theta}))$, holds at the optimum.*

Proof. Take any $(y(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}\in C^2[\underline{\theta},\bar{\theta}]$ such that $y(\bar{\theta})=0$ and note that there exists $\varepsilon>0$ such that $(v^*(\theta)+\lambda y(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$ satisfies the monotonicity constraint for $|\lambda|<\varepsilon$. If we let

$$g(\lambda)=\int_{\underline{\theta}}^{\bar{\theta}}\Lambda(\theta,v^*(\theta)+\lambda y(\theta),\dot{v}^*(\theta)+\lambda\dot{y}(\theta))f(\theta)d\theta,$$

a necessary condition for optimality is $g'(0)=0$. Therefore, the lemma follows from Theorem 14.19 in Clarke (2013). □

Lemma 7. *If χ implements the planner's solution then the functional $M_\theta(z)$ defined by $M_\theta(z):=\Lambda(\theta,v^*(\theta),z)$ satisfies $M'_\theta(\dot{v}^*(\theta))\leq 0$ for almost all θ.*

Proof. From Lemma 5, Λ is a smooth function. Notice that for any C^2 function, $(y(\theta))_{\theta\in[\underline{\theta},\bar{\theta}]}$, there exists $\varepsilon>0$ such that $|\lambda|<\varepsilon$ implies $\dot{v}^*(\theta)+\lambda\dot{y}(\theta)<0$ for all θ. Hence $v^*(\theta)+\lambda y(\theta)$ is monotonic. The result follows from the argument from Theorem 14.7 in Clarke (2013). □

D Bunching

D.1 Problems with Bounded Constraint

Consider the program, $\mathcal{P}_n$,

$$\max \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n(\theta)) - \kappa\left(v_n(\theta) + \theta h_n(\theta)\right)\right] f(\theta) d\theta$$

Subject to

$$h_n(\bar{\theta}) \geq 0 \text{ , } h_n(\underline{\theta}) \leq H,$$
$$-\dot{h}_n(\bar{\theta}) \in \left[n^{-1}, n\right],$$

and

$$\int_{\underline{\theta}}^{\bar{\theta}} v(\theta) f(\theta) d(\theta) \geq v_0.$$

For any allocation, $(u_n(\theta), h_n(\theta))_{\theta\in\Theta}$, let

$$\mathcal{V}\left[(u_n(\theta), h_n(\theta))_{\theta\in\Theta}\right] := \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n(\theta)) - \kappa\left(v_n(\theta) + \theta h_n(\theta)\right)\right] f(\theta) d\theta.$$

Let $(u_n^*(\theta), h_n^*(\theta))_{\theta\in\Theta}$ be a solution to $\mathcal{P}_n$.

Lemma 8. *Problem $\mathcal{P}_n$ admits a solution. Moreover, all its solutions are [a.e.] equal.*

Proof. The assumption that $v_0 \geq 0$ implies that the allocation given by $u_n(\theta) = v_0\ \forall\theta$ and $h_n(\theta) = 0\ \forall\theta$ is feasible $\forall n$. The existence of a solution then follows from Tonelli Theorem (Theorem 16.2 in Clarke (2013)).[16]

To see why the solution must be unique, consider two candidate solutions, $(u_n^a(\theta), h_n^a(\theta))_{\theta\in\Theta}$, and $\left(u_n^b(\theta), h_n^b(\theta)\right)_{\theta\in\Theta}$. For $\lambda \in (0,1)$, let $\left(u_n^\lambda(\theta), h_n^\lambda(\theta)\right)_{\theta\in\Theta}$ be defined by $u_n^\lambda(\theta) = \lambda u_n^a(\theta) + (1-\lambda) u_n^b(\theta)\ \forall\theta$ and $h_n^\lambda(\theta) = \lambda h_n^a(\theta) + (1-\lambda) h_n^b(\theta)\ \forall\theta$. Note that

$$u_n^a(\theta) - \theta h_n^a(\theta) \geq u_n^a(\hat{\theta}) - \theta h_n^a(\hat{\theta}),$$

and

$$u_n^b(\theta) - \theta h_n^b(\theta) \geq u_n^b(\hat{\theta}) - \theta h_n^b(\hat{\theta}),$$

implies

$$u_n^\lambda(\theta) - \theta h_n^\lambda(\theta) \geq u_n^\lambda(\hat{\theta}) - \theta h_n^\lambda(\hat{\theta}).$$

Next, for $i = a, b$, let $v_n^i(\theta) := u_n^i(\theta) - \theta h_n^i(\theta)$ and notice that

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\lambda v_n^a(\theta) + (1-\lambda) v_n^b(\theta)\right] f(\theta) d(\theta) \geq$$
$$\min\left\{\int_{\underline{\theta}}^{\bar{\theta}} v_n^a(\theta) f(\theta) d(\theta), \int_{\underline{\theta}}^{\bar{\theta}} v_n^b(\theta) f(\theta) d(\theta)\right\} \geq v_0.$$

[16] Theorem 16.2 does not assume that $\dot{h}_n(\bar{\theta}) \in \left[n^{-1}, n\right]$. The changes required to handle this case are straightforward – see proof of Lemma 3.

Thus, from the strict concavity of the objective function we have

$$\mathcal{V}\left[\left(u_n^{\lambda}(\theta), h_n^{\lambda}(\theta)\right)_{\theta\in\Theta}\right] > \min\left\{\mathcal{V}\left[\left(u_n^{a}(\theta), h_n^{a}(\theta)\right)_{\theta\in\Theta}\right], \mathcal{V}\left[\left(u_n^{b}(\theta), h_n^{b}(\theta)\right)_{\theta\in\Theta}\right]\right\}.$$

Because the $\left(u_n^{\lambda}(\theta), h_n^{\lambda}(\theta)\right)_{\theta\in\Theta}$ is incentive compatible, it cannot be true that both $\left(u_n^{a}(\theta), h_n^{a}(\theta)\right)_{\theta\in\Theta}$, and $\left(u_n^{b}(\theta), h_n^{b}(\theta)\right)_{\theta\in\Theta}$ solve $\mathcal{P}_n$. □

Lemma 9. *The sequence* $\left\{\mathcal{V}\left[\left(u_n^{*}(\theta), h_n^{*}(\theta)\right)_{\theta\in\Theta}\right]\right\}_n$ *converges to* $\mathcal{V}\left[\left(u^{*}(\theta), h^{*}(\theta)\right)_{\theta\in\Theta}\right]$.

Proof. Towards a contradiction, assume that there is $\varepsilon > 0$ such that

$$\liminf \mathcal{V}\left[\left(u_n^{*}(\theta), h_n^{*}(\theta)\right)_{\theta\in\Theta}\right] < \mathcal{V}\left[\left(u^{*}(\theta), h^{*}(\theta)\right)_{\theta\in\Theta}\right] - \varepsilon. \tag{30}$$

For η sufficiently small, the allocation

$$\left(u_\eta^{*}(\theta), h_\eta^{*}(\theta)\right)_{\theta\in\Theta} := \left(u^{*}(\theta), h^{*}(\theta) - \eta\right)_{\theta\in\Theta}$$

is also feasible and yields a payoff of at least $\mathcal{V}\left[\left(u^{*}(\theta), h^{*}(\theta)\right)_{\theta\in\Theta}\right] - \frac{\varepsilon}{2}$. Moreover, this allocation leads to a utility profile $\left(v_\eta^{*}(\theta)\right)_{\theta\in\Theta}$ that satisfies $\int_{\underline{\theta}}^{\bar{\theta}} v_\eta^{*}(\theta) f(\theta) d(\theta) \geq v_0 + \eta\mathbb{E}[\theta]$.

Notice that $h_\eta^{*}(\cdot)$ is bounded and decreasing. Therefore, for each $m \in \mathbb{N}$ there exists a decreasing and C^2 function $h_\eta^{m}(\cdot)$ such that

$$\int_{\underline{\theta}}^{\bar{\theta}} \left|h_\eta^{*}(\theta) - h_\eta^{m}(\theta)\right| f(\theta) d\theta < m^{-1},$$

and $h_\eta^{m}(\bar{\theta}) = h_\eta^{*}(\bar{\theta})$.

Construct the allocation $\left(u_\eta^{m}(\theta), h_\eta^{m}(\theta)\right)_{\theta\in\Theta}$ from $h_\eta^{m}(\cdot)$ through $u_\eta^{m}\left(\bar{\theta}\right) = u^{*}\left(\bar{\theta}\right)$ and $\dot{u}_\eta^{m}(\theta) = \dot{h}_\eta^{m}(\theta)$. The allocation is trivially implementable. Letting $\left(v_\eta^{m}(\theta)\right)_{\theta\in\Theta}$ be the utility profile associated with $\left(u_\eta^{m}(\theta), h_\eta^{m}(\theta)\right)_{\theta\in\Theta}$, we have

$$\left|v_\eta^{m}(\theta) - v_\eta^{*}(\theta)\right| = \int_{\theta}^{\bar{\theta}} \left|h_\eta^{*}(a) - h_\eta^{m}(a)\right| f(a) da \leq m^{-1}. \tag{31}$$

From (31), there exists $m^* \in \mathbb{N}$ such that $\int_{\underline{\theta}}^{\bar{\theta}} v_\eta^{m}(\theta) f(\theta) d(\theta) \geq v_0$ for all $m > m^*$. From (31) and

$$\left|u_\eta^{m}(\theta) - u_\eta^{*}(\theta)\right| = \left|v_\eta^{m}(\theta) + \theta h_\eta^{m} - \left(v_\eta^{*}(\theta) + \theta h_\eta^{m}\right)\right|,$$

we have

$$\int_{\theta}^{\bar{\theta}} \left|u_\eta^{m}(a) - u_\eta^{*}(a)\right| f(a) da \leq m^{-1} + \int_{\theta}^{\bar{\theta}} \left|\theta\left(h_\eta^{*}(a) - h_\eta^{m}(a)\right)\right| f(a) da$$
$$\leq m^{-1} + \bar{\theta}\int_{\theta}^{\bar{\theta}} \left|h_\eta^{*}(a) - h_\eta^{m}(a)\right| f(a) da = \left(\frac{1+\bar{\theta}}{m}\right).$$

Hence, we can take a subsequence of $\left(u_\eta^{m}(\theta), h_\eta^{m}(\theta)\right)_{\theta\in\Theta}$ which converges almost everywhere to $\left(u_\eta^{*}(\theta), h_\eta^{*}(\theta)\right)_{\theta\in\Theta}$. Since h lies in a bounded set, we conclude that

$\left\{\left(u_\eta^m(\theta), h_\eta^m(\theta)\right)_{\theta\in\Theta}\right\}_{m=1}^{\infty}$ as well as $\left(u_\eta^*(\theta), h_\eta^*(\theta)\right)_{\theta\in\Theta}$ are uniformly bounded. ϕ and κ are continuous functions so we can apply Lebesgue's Dominated Convergence Theorem to conclude that

$$\lim \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_\eta^m(\theta)) - \kappa\left(u_\eta^m(\theta)\right)\right] f(\theta)d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_\eta^*(\theta)) - \kappa\left(u_\eta^*(\theta)\right)\right] f(\theta)d\theta.$$

This immediately implies that there is $m^{**} \geq m^*$ such that

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_\eta^{m^{**}}(\theta)) - \kappa\left(u_\eta^{m^{**}}(\theta)\right)\right] f(\theta)d\theta \geq \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_\eta^*(\theta)) - \kappa\left(u_\eta^*(\theta)\right)\right] f(\theta)d\theta - \frac{\varepsilon}{4}, \quad (32)$$

which implies that

$$\mathcal{V}\left[\left(u_\eta^{m^{**}}(\theta), h_\eta^{m^{**}}(\theta)\right)_{\theta\in\Theta}\right] \geq \mathcal{V}\left[\left(u^*(\theta), h^*(\theta)\right)_{\theta\in\Theta}\right] - \frac{3\varepsilon}{4}.$$

Since $\left(u_\eta^{m^{**}}(\theta), h_\eta^{m^{**}}(\theta)\right)_{\theta\in\Theta}$ is feasible for the problem $\mathcal{P}_n$ whenever n is large enough, this contradicts (30). □

Lemma 10. *There exists a subsequence of solutions, $\left(u_n^*(\theta), h_n^*(\theta)\right)_{\theta\in\Theta}$, which converges almost everywhere to the allocation $\left(u^*(\theta), h^*(\theta)\right)_{\theta\in\Theta}$.*

Proof. Since both mappings, u_n^* and h_n^*, are monotonic, Helly's Selection Theorem implies that we can find a subsequence which [a.e.] converges to the allocation $\left(u^\&(\theta), h^\&(\theta)\right)_{\theta\in\Theta}$. In particular, the subsequence $\left(u_n^*(\theta), h_n^*(\theta)\right)$ (without relabelling) converges to $\left(u^\&(\theta), h^\&(\theta)\right)$ in all points of continuity of $\left(u^\&(\theta), h^\&(\theta)\right)$.

We can assume, without loss, that $\left(u^\&(\theta), h^\&(\theta)\right)_{\theta\in\Theta}$ is right-continuous. To prove that $\left(u^\&(\theta), h^\&(\theta)\right)_{\theta\in\Theta}$ is implementable, it suffices to show that for any point of continuity of $\left(u^\&(\theta), h^\&(\theta)\right)_{\theta\in\Theta}$ we have

$$u^\&(\theta) - \theta h^\&(\theta) \geq u^\&(\theta') - \theta h^\&(\theta')$$

for all θ'.[17]

From the right-continuity of the limit allocation, we have

$$u^\&(\theta') - \theta h^\&(\theta') = \lim u^\&(\theta_k) - \theta h^\&(\theta_k),$$

where $(\theta_k) \downarrow \theta$ is a sequence of points of continuity of $\left(u^\&(\theta), h^\&(\theta)\right)_{\theta\in\Theta}$. Fix θ_k and notice that since $\left(u_n(\theta), h_n(\theta)\right)_{\theta\in\Theta}$ is implementable we have, for all n, $u_n(\theta) - \theta h_n(\theta) \geq u_n(\theta_k) - \theta h_n(\theta_k)$, which implies

$$u^\&(\theta) - \theta h^\&(\theta) = \lim u_n(\theta) - \theta h_n(\theta) \geq \lim u_n(\theta_k) - \theta h_n(\theta_k) = u^\&(\theta_k) - \theta h^\&(\theta_k).$$

[17]Since $\left(u_n^\&(\theta), h_n^\&(\theta)\right)_{\theta\in\Theta}$ is monotonic its points of discontinuities lie in a countable set.

Next, notice that we have

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^{\&}(\theta)) - \kappa\left(u^{\&}(\theta)\right)\right] f(\theta)d\theta = \lim \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n(\theta)) - \kappa\left(u_n(\theta)\right)\right] f(\theta)d\theta,$$

by the Lebesgue Dominated Theorem. From Lemma 9, $\left(u^{\&}(\theta), h^{\&}(\theta)\right)_{\theta\in\Theta}$ is optimal, while from Lemma 8,

$$\left(u^{\&}(\theta), h^{\&}(\theta)\right) = \left(u^{*}(\theta), h^{*}(\theta)\right) a.e.$$

□

The following Lemma is a corollary from Lemma 10.

Lemma 11. *For any* $\varepsilon > 0$, *there exists* $n_1 \in \mathbb{N}$ *such that* $n \geq n_1$ *implies* $u_n^*(\theta) \in \left[u^*\left(\bar{\theta}\right) - \varepsilon, u^*\left(\bar{\theta}\right) + \varepsilon\right]$ $\forall\theta$ *and* $h_n^*(\theta) \in \left[h^*\left(\bar{\theta}\right) - \varepsilon, h^*\left(\bar{\theta}\right) + \varepsilon\right]$ $\forall\theta$.

For $\varepsilon > 0$ sufficiently small assume, without loss, that $u_n^*(\theta) \in \left[u^*\left(\bar{\theta}\right) - \varepsilon, u^*\left(\underline{\theta}\right) + \varepsilon\right]$ $\forall\theta$ and $h_n^*(\theta) \in \left[h^*\left(\bar{\theta}\right) - \varepsilon, h^*\left(\underline{\theta}\right) + \varepsilon\right]$ $\forall\theta$ (otherwise take the tail subsequence starting at n_1).

D.2 Characterizing the Sequences of Programs

D.2.1 Problem $\mathcal{P}_n$

The planner's program, $\mathcal{P}_n$, has the following optimal-control formulation,

$$\begin{aligned}\mathcal{H}_n = \left[\phi(h_n(\theta)) - \kappa\left(v_n(\theta) + \theta h_n(\theta)\right)\right] f(\theta) + \gamma_n(\theta)\left[\dot{\Phi}_n(\theta) - v_n(\theta)f(\theta)\right] \\ + \lambda_n(\theta)\left[\dot{h}_n(\theta) + x_n(\theta)\right] + \mu_n(\theta)\left[\dot{v}_n(\theta) + h_n(\theta)\right] \\ + \underline{\alpha}_n(\theta)\left[x_n(\theta) - n^{-1}\right] + \overline{\alpha}_n(\theta)\left[n - x_n(\theta)\right]\end{aligned}$$

The state variables are h_n ,v_n and Φ_n, while the control variable is x_n. In this problem, x_n is restricted to lie in $[n^{-1}, n]$.

We have the following boundary conditions,

(i) $h_n(\underline{\theta})$ and $h_n\left(\bar{\theta}\right)$ free.

(ii) $v_n(\underline{\theta})$ and $v_n\left(\bar{\theta}\right)$ free.

(iii) $\dot{\Phi}_n(\underline{\theta}) = 0$ and $\dot{\Phi}_n(\bar{\theta}) = v_0$.

The optimality conditions are, in this case,

$$\lambda_n(\theta) + \underline{\alpha}_n(\theta) - \overline{\alpha}_n(\theta) = 0, \tag{33}$$

$$-\dot{\lambda}_n(\theta) = \left[\phi'(h_n(\theta)) - \theta\kappa'\left(v_n(\theta) + \theta h_n(\theta)\right)\right] f(\theta) + \mu_n(\theta), \tag{34}$$

$$-\dot{\mu}_n(\theta) = -\kappa'\left(v_n(\theta) + \theta h_n(\theta)\right) f(\theta) - \gamma_n(\theta) f(\theta), \tag{35}$$

and

$$-\dot{\gamma}_n(\theta) = 0. \tag{36}$$

First, note from (36) that γ_n is a constant. Next, integrating $\dot{\mu}_n(\theta)$ in (35) we obtain

$$\mu_n(\theta) = \int_{\underline{\theta}}^{\theta} \left[\kappa'(v_n(a) + a h_n(a)) - \gamma_n\right] f(a) da.$$

Using the fact that $h_n(\bar{\theta})$ is free, we have from (35)

$$\gamma_n = \int_{\underline{\theta}}^{\bar{\theta}} \kappa'(v_n(\theta) + \theta h_n(\theta)) f(\theta) d\theta.$$

Using the expression for $-\dot{\lambda}_n(\theta)$, we have:

$$-\dot{\lambda}_n(\theta) = \left[\phi'(h_n(\theta)) - \theta\kappa'(v_n(\theta) + \theta h_n(\theta))\right] f(\theta) + \int_{\underline{\theta}}^{\theta} \left[\kappa'(v_n(a) + a h_n(a)) - \gamma_n\right] f(a) da.$$

Since $v_n(\underline{\theta})$ and $v_n(\bar{\theta})$ are free, we have $\lambda_n(\underline{\theta}) = \lambda_n(\bar{\theta}) = 0$, thus, $\lambda_n(\theta) = -\int_{\theta}^{\bar{\theta}} \dot{\lambda}_n(a) da$. In this case,

$$\lambda_n(\theta) = \int_{\underline{\theta}}^{\theta} \Bigg[\phi'\left(h_n(\tilde{\theta})\right) - \theta\kappa'\left(v_n\left(\tilde{\theta}\right) + \tilde{\theta} h_n(\tilde{\theta})\right) f(\tilde{\theta}) + \int_{\underline{\theta}}^{\tilde{\theta}} \left[\kappa'(v_n(a) + a h_n(a)) - \gamma_n\right] f(a) da\Bigg] d\tilde{\theta}.$$

Moreover, notice that $x_n(\theta) = n^{-1} \Rightarrow \underline{\alpha}_n(\theta) \geq 0 \Rightarrow \lambda_n(\theta) \leq 0$, $x_n(\theta) = n \Rightarrow \bar{\alpha}_n \geq 0 \Rightarrow \lambda_n(\theta) \geq 0$ and $x_n(\theta) \in (n^{-1}, n) \Rightarrow \lambda_n(\theta) = 0$.

Now we analyze the firm's program.

D.2.2 Problem $\mathcal{P}_n^F$

The firm's program, $\mathcal{P}_n^F$, has the following optimal-control formulation,

$$\mathcal{H}_n = \left[\phi(h_n(\theta)) - \chi_n(v_n(\theta) + \theta h_n(\theta))\right] f(\theta) + \lambda_n^F(\theta)\left[\dot{h}_n(\theta) + x_n(\theta)\right] + \mu_n^F(\theta)\left[\dot{v}_n(\theta) + h_n(\theta)\right] + \underline{\alpha}_n^F(\theta)\left[x_n(\theta) - n^{-1}\right] + \bar{\alpha}_n^F(\theta)\left[n - x_n(\theta)\right]$$

The state variables are h_n and v_n, and the control variable is x_n. In this problem, x_n is restricted to lie in $[n^{-1}, n]$. Furthermore, we define χ_n for the domain $\left[u^*\left(\bar{\theta}\right) - \varepsilon, u^*\left(\bar{\theta}\right) + \varepsilon\right]$. We may set $\chi_n(u) = \infty \; \forall u \in \left[u^*\left(\bar{\theta}\right) - \varepsilon, u^*\left(\bar{\theta}\right) + \varepsilon\right]^c$.

We have the following boundary conditions:

(i) $h_n(\underline{\theta})$ and $h_n(\bar{\theta})$ free;

(ii) $v_n(\underline{\theta})$ free, and;

(iii) $v_n(\bar{\theta}) = v_n^*(\bar{\theta})$.

The following are the optimality conditions,

$$\lambda_n^F(\theta) + \underline{\alpha}_n^F(\theta) - \overline{\alpha}_n^F(\theta) = 0, \tag{37}$$

$$-\dot{\lambda}_n^F(\theta) = \left[\phi'(h_n(\theta)) - \theta\chi_n'(v_n(\theta) + \theta h_n(\theta))\right] f(\theta) + \mu_n^F(\theta), \tag{38}$$

and

$$-\dot{\mu}_n^F(\theta) = -\chi_n'(v_n(\theta) + \theta h_n(\theta)) f(\theta). \tag{39}$$

D.2.3 Connecting $\mathcal{P}_n^F$ with $\mathcal{P}_n$

We will construct $\left(\lambda_n^F(\theta), \mu_n^F(\theta), \chi_n(\theta), \underline{\alpha}_n^F(\theta), \bar{\alpha}_n^F(\theta)\right)$ to guarantee that all necessary conditions of the firm's problem are satisfied when they are evaluated at $(u_n^*(\theta), h_n^*(\theta))$. Moreover, we must pick $\chi_n(\cdot)$ in such a way that the firm's problem is concave. Once we have accomplished this, Theorem 24.1 in Clarke (2013) guarantees that $(u_n^*(\theta), h_n^*(\theta))$ solves the firm's problem.

Using the fact that $v_n(\underline{\theta})$ is free, we define $\mu_n^F(\underline{\theta}) := 0$, which implies $\mu_n^F(\theta) = \int_{\underline{\theta}}^{\theta} \dot{\mu}_n^F(a)\, da$. Hence, from (38) and (39),

$$-\dot{\lambda}_n^F(\theta) = \left[\phi'(h_n(\theta)) - \theta\chi_n'(v_n(\theta) + \theta h_n(\theta))\right] f(\theta) + \int_{\underline{\theta}}^{\theta} \chi_n'(v_n(a) + a h_n(a)) f(a) da.$$

Next, define $\chi_n(\cdot)$ such that $\chi_n'(v_n^*(\underline{\theta}) + \theta h_n^*(\underline{\theta})) = \kappa'(v_n^*(\underline{\theta}) + \theta h_n^*(\underline{\theta}))$. Define also $\left(\underline{\alpha}_n^F(\theta), \overline{\alpha}_n^F(\theta)\right) := (\underline{\alpha}_n(\theta), \overline{\alpha}_n(\theta))$ and $\lambda_n^F(\underline{\theta}) := \lambda_n(\underline{\theta}) = 0$. We will construct $\chi_n(\cdot)$ in such a way that

$$\dot{\lambda}_n^F(\theta) = \dot{\lambda}_n(\theta) \;\; \forall \theta. \tag{40}$$

Given $\lambda_n^F(\underline{\theta}) = \lambda_n(\underline{\theta})$, (40) this implies that $\lambda_n(\theta) = \lambda_n(\theta)$ for all θ and, in particular, that $\lambda_n^F(\bar{\theta}) = \lambda_n(\bar{\theta})$. Taking all this into account we have,

$$\lambda_n^F(\theta) + \underline{\alpha}_n^F(\theta) - \overline{\alpha}_n^F(\theta) = \lambda_n(\theta) + \underline{\alpha}(\theta) - \overline{\alpha}_n(\theta) = 0$$

and $x_n(\theta) = n^{-1} \Rightarrow \underline{\alpha}_n^F(\theta) \geq 0 \Rightarrow \lambda_n^F(\theta) \leq 0$, $x_n(\theta) = n \Rightarrow \overline{\alpha}_n^F(\theta) \geq 0 \Rightarrow \lambda_n^F(\theta) \geq 0$ and $x_n(\theta) \in (n^{-1}, n) \Rightarrow \lambda_n^F(\theta) = 0$. Therefore, all optimality conditions hold.

To obtain (40) we must show that for all θ we have

$$\left[\phi'(h_n^*(\theta)) - \theta\chi_n'(v_n^*(\theta) + \theta h_n^*(\theta))\right] f(\theta) + \int_{\underline{\theta}}^{\theta} \chi_n'(v_n^*(a) + a h_n^*(a)) f(a) da = \left[\phi'(h_n^*(\theta)) - \theta\kappa'(v_n^*(\theta) + \theta h_n^*(\theta))\right] f(\theta) + \int_{\underline{\theta}}^{\theta} \left[\kappa'(v_n^*(a) + a h_n^*(a)) - \gamma_n\right] f(a) da. \tag{41}$$

Since (41) holds by construction at $\underline{\theta}$, it suffices to show that

$$\theta\left[\chi_n'(v_n^*(\theta) + \theta h_n^*(\theta)) - \kappa'(v_n^*(\theta) + \theta h_n^*(\theta))\right] f(\theta) = \int_{\underline{\theta}}^{\theta} \left[\chi_n'(v_n^*(a) + a h_n^*(a)) - \kappa'(v_n^*(\theta) + \theta h_n^*(\theta))\right] f(a) da.$$

Defining $\rho(\theta) := \left[\chi_n'\left(v_n^*(\theta) + \theta h_n^*(\theta)\right) - \kappa'\left(v_n^*(\theta) + \theta h_n^*(\theta)\right)\right] f(\theta)$, we must guarantee that

$$\theta\rho(\theta) = -\int_{\underline{\theta}}^{\theta} \rho(a)\, da \;\forall\theta.$$

This equation has a unique solution subject to $\rho(\underline{\theta}) = 0$. We can, therefore, pin down χ_n uniquelly. The convexity of $\chi_n'\left(v_n^*(\theta) + \theta h_n^*(\theta)\right)$ follows from our assumption on F, from the fact that $-\dot{h}_n^*(\theta) \in [n^{-1}, n]$, and from

$$\chi_n''\left(v_n^*(\theta) + \theta h_n^*(\theta)\right) = \kappa''\left(v_n^*(\theta) + \theta h_n^*(\theta)\right) - \frac{\rho'(\theta)}{2\theta\dot{h}_n^*(\theta)}.$$

Finally, we extend χ_n continuously to $\left[u^*\left(\bar{\theta}\right) - \varepsilon, u^*\left(\underline{\theta}\right) + \varepsilon\right]$ by making χ_n' constant in $\left[u^*\left(\bar{\theta}\right) - \varepsilon, u_n^*\left(\bar{\theta}\right)\right]$ and in $\left[u_n^*\left(\underline{\theta}\right), u^*\left(\underline{\theta}\right) + \varepsilon\right]$.

We have implicitly assumed that the firm never excludes any type. The following Lemma is straightforward and its proof is omitted.

Lemma 12. *There exists $\underline{\chi} \in \mathbb{R}$ such that if $\chi_n\left(u^*\left(\underline{\theta}\right) + \varepsilon\right) = \underline{\chi}$ for all n, then it is optimal for the firm to hire all types in every program $\mathcal{P}_n^F$.*

We assume that $\chi_n\left(u^*\left(\underline{\theta}\right) + \varepsilon\right) = \underline{\chi}$ for all n for the remainder of this proof.

D.3 Constructing $\chi(\cdot)$ to support $\left(u^*(\theta), h^*(\theta)\right)$ in the unrestricted firm's problem

Lemma 13. *There exists a subsequence of (χ_n) which converges uniformly to a Lipschitz mapping χ.*

Proof. Notice that

$$\sup_n \max_{u \in \left[u^*\left(\bar{\theta}\right) - \varepsilon, u_n^*\left(\underline{\theta}\right) + \varepsilon\right]} \left|\chi_n'(u)\right| \leq \max_\theta \rho(\theta) + \max_{u \in \left[u^*\left(\bar{\theta}\right) - \varepsilon, u_n^*\left(\underline{\theta}\right) + \varepsilon\right]} \kappa'(u) < \infty.$$

Hence,

$$\sup_n \max_{u \in \left[u^*\left(\bar{\theta}\right) - \varepsilon, u_n^*\left(\bar{\theta}\right)\right]} \left|\chi_n(u)\right| \leq$$

$$\underline{\chi} + \left[2\varepsilon + \left(u_n^*\left(\underline{\theta}\right) - u^*\left(\bar{\theta}\right)\right)\right] \left(\max_\theta \rho(\theta) + \max_{u \in \left[u^*\left(\bar{\theta}\right) - \varepsilon, u_n^*\left(\underline{\theta}\right) + \varepsilon\right]} \kappa'(u)\right) < \infty.$$

The result, then, follows from the Arzelà-Ascoli theorem. □

Next, consider the firm's unrestricted problem. The firm must implementable allocations $\left(u(\theta), h(\theta)\right)_{\theta \in \Theta}$ which deliver utility $v^*(\bar{\theta})$ to type $\bar{\theta}$.

Proposition 8. *The allocation $\left(u^*(\theta), h^*(\theta)\right)$ solves the firm's problem for the tax-function χ.*

Proof. Assume towards a contradiction that there exists $\eta > 0$ and an implementable allocation $\left(u^{\#}(\theta), h^{\#}(\theta)\right)_{\theta\in\Theta}$ which yields the indirected utility $v^*(\bar{\theta})$ to type $\bar{\theta}$ and for which we have

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^{\#}(\theta)) - \chi\left(u^{\#}(\theta)\right)\right] f(\theta)d\theta > \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^{*}(\theta)) - \chi\left(u^{*}(\theta)\right)\right] f(\theta)d\theta + \eta.$$

It is easy to show that we can find $m \in \mathbb{N}$ and an allocation $\left(u_m^{\#}(\theta), h_m^{\#}(\theta)\right)_{\theta\in\Theta}$ such that

a) $h_m^{\#}(\theta)$ is absolutely continuous and $\dot{h}_m^{\#}(\theta) \in [m^{-1}, m]$;

b) $v_m^{\#}(\bar{\theta}) \geq v^*(\bar{\theta}) + m^{-1}$,

c)

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_m^{\#}(\theta)) - \chi\left(u_m^{\#}(\theta)\right)\right] f(\theta)d\theta > \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^{*}(\theta)) - \chi\left(u^{*}(\theta)\right)\right] f(\theta)d\theta + \frac{3\eta}{4}.$$

Next, notice that there exists $n_1 \in \mathbb{N}$ such that $\left(u_m^{\#}(\theta), h_m^{\#}(\theta)\right)_{\theta\in\Theta}$ is feasible for problem $\mathcal{P}_n^F$ for all $n \geq n_1$. Thus, by the Lebesgue Dominated Convergence theorem

$$\lim \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_m^{\#}(\theta)) - \chi_n\left(u_m^{\#}(\theta)\right)\right] f(\theta)d\theta = \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_m^{\#}(\theta)) - \chi\left(u_m^{\#}(\theta)\right)\right] f(\theta)d\theta,$$

which implies that there exists $n_2 \geq n_1$ such that, for all $n \geq n_2$,

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_m^{\#}(\theta)) - \chi_n\left(u_m^{\#}(\theta)\right)\right] f(\theta)d\theta \geq \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^{*}(\theta)) - \chi\left(u^{*}(\theta)\right)\right] f(\theta)d\theta + \frac{\eta}{2}. \quad \textbf{(42)}$$

Next, notice that since $\left(u_n^*(\theta), h_n^*(\theta)\right)$ converges almost everywhere to $\left(u^*(\theta), h^*(\theta)\right)_{\theta\in\Theta}$ we have (again by the Lebesgue Dominated Convergence theorem):

$$\lim \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n^{*}(\theta)) - \chi\left(u_n^{*}(\theta)\right)\right] f(\theta)d\theta = \lim \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^{*}(\theta)) - \chi\left(u^{*}(\theta)\right)\right] f(\theta)d\theta,$$

which implies that there exists $n_3 \geq n_2$ such that, for all $n \geq n_3$,

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n^{*}(\theta)) - \chi\left(u_n^{*}(\theta)\right)\right] f(\theta)d\theta < \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^{*}(\theta)) - \chi\left(u^{*}(\theta)\right)\right] f(\theta)d\theta + \frac{\eta}{8}. \quad \textbf{(43)}$$

On the other hand,

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n^*(\theta)) - \chi_n\left(u_n^*\left(\theta\right)\right)\right] f(\theta) d\theta =$$

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n^*(\theta)) - \chi\left(u_n^*\left(\theta\right)\right)\right] f(\theta) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} \left[\chi_n\left(u_n^*\left(\theta\right)\right) - \chi\left(u_n^*\left(\theta\right)\right)\right] f(\theta) d\theta$$

$$\leq \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n^*(\theta)) - \chi\left(u_n^*\left(\theta\right)\right)\right] f(\theta) d\theta + \left\|\chi_n - \chi\right\|. \quad \textbf{(44)}$$

From Lemma 13 there exists $n_4 \geq n_3$ such that $n \geq n_4$ implies $\|\chi_n - \chi\| \leq \frac{\eta}{8}$. Thus for all $n \geq n_4$, using (42), (43) and (44) we have:

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_m^{\#}(\theta)) - \chi_n\left(u_m^{\#}\left(\theta\right)\right)\right] f(\theta) d\theta \geq \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h^*(\theta)) - \chi\left(u^*\left(\theta\right)\right)\right] f(\theta) d\theta + \frac{\eta}{2}$$

$$> \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n^*(\theta)) - \chi\left(u_n^*\left(\theta\right)\right)\right] f(\theta) d\theta + \frac{3\eta}{8}$$

$$\geq \int_{\underline{\theta}}^{\bar{\theta}} \left[\phi(h_n^*(\theta)) - \chi_n\left(u_n^*\left(\theta\right)\right)\right] f(\theta) d\theta + \frac{\eta}{4},$$

which contradicts the optimality of $\left(u_n^*\left(\theta\right), h_n^*(\theta)\right)$ for the problem $\mathcal{P}_n^F$. □