

Redistribution with Labor Markets Frictions*

Carlos E. da Costa[†] Lucas J. Maestri[‡] Marcelo R. Santos[§]

VERY, VERY PRELIMINARY

Abstract

How does the presence of labor market frictions change optimal redistributive policies? Embedding an otherwise standard Mirrlees' (1971) economy in a directed search environment we show that optimal Utilitarian (and Rawlsian) allocations are characterized by downward distortions in labor supply in *both* the intensive and extensive margins. After showing that income tax schedules alone cannot implement all constrained efficient allocations, we quantify, for the US economy, the costs of relying on optimal labor income taxes alone.

Keywords: *Mirrlees' problem; Directed Search.* **JEL Classification:** *D82, H21.*

[$\text{raise}=0.1, \text{nindent}=0\text{em}, \text{slope}=-.5\text{em}$]Optimal redistributive policy in the [Mirrlees' \(1971\)](#) tradition normally assumes that labor markets are frictionless. However useful as a first approximation, this assumption precludes the analysis of, among other things, the consequences for unemployment of tax design.

In this paper we consider that trade in labor market is decentralized and frictional. As in [Goloso et al. \(2013\)](#) frictions arise due to our assuming that the matching function which summarizes the process which brings together firms and workers interested in the same labor contracts, does not guarantee that a worker will find a job nor that a vacancy will be filled. In contrast with [Goloso et al. \(2013\)](#) we assume that heterogeneity is on the side of workers while firms are identical. Hence, we do not capture the type of ex-post wage dispersion and residual inequality that motivates their work. We focus instead on the consequences of labor

*da Costa thanks CNPq project 301140/2017-0 for financial support. All errors are our responsibility.

[†]EPGE FGV Brazilian School of Economics and Finance

[‡]EPGE FGV Brazilian School of Economics and Finance

[§]Inspere

market frictions for optimal *redistributive* policies. Moreover, the intensive margin of labor supply, which is absent from most work that takes unemployment into account, is also examined here.

Posting a vacancy is costly for the firms. Once a vacancy is posted, the probability that it is filled depends on labor market tightness. Unemployment arises since the assumptions we make about the matching function leads to a less than unit probability that a worker gets a job even if he or she is willing to work at the posted labor contract. Workers then apply to vacancies that maximize their expected utility.

In this environment, a complete characterization of optimal Utilitarian and Rawlsian policies is provided. We show that at the optimum both effort (intensive margin) and participation (extensive margin) are increasing in agents' productivity. Note that incentive compatibility in this setting only implies that total hours, the combination of extensive and intensive margins, be increasing in productivity. It is optimality which implies that *both* margins are. We use this result to prove that both effort conditional on being employed and employment are lower than in the first best, thus extending the scope of [Mirrlees' \(1971\)](#) findings.

By following [Mirrlees'](#) mechanism design approach, we, characterize constrained efficient allocations under the assumption that the planner chooses, earnings, taxes and the number of offers made by each firm. We then ask whether a informationally feasible unemployment benefit along with a non-linear income tax schedule suffice to implement the optimal allocation. We find it *not* to be the case, and show how tax schedules that use information on the revenues generated by the firm to which an agent works can implement the optimum.

Because such taxes are seldom used, and not in the US, we approximate the current US schedule with the parametric function of [Musgrave \(1959\)](#); [Bénabou \(2000, 2002\)](#) to provide a quantitative assessment of the potential losses from relying on unemployment benefits and labor income taxes only.

The rest of the paper is organized as follows. In Section 1 we describe the model economy. We describe the Planner's program and characterize its solution in Section 2. In Section 3 we show that a non-linear income tax schedule alone cannot implement the constrained efficient allocation. Finally, in Section 4 we assess the cost for the US government of not having (or not using) richer instruments capable of manipulating both extensive and intensive margin choices.

Literature Review

In [Mirrlees' \(1971\)](#) seminal work, a characterization of constrained efficient allocations is offered in a world where agents with private information about their productivities must be offered incentives for truthfully revealing them. This mechanism design approach to redistribution policies which has been dominant until the end of the 20th century was often criticized for its oversimplified environment which precluded the discussion of important aspects of labor market policies. Among them, is the nonexistence of involuntary unemployment in [Mirrlees' \(1971\)](#) model.¹ In this paper we try to fill this gap.

Ours is not the first work to consider taxation in the presence of labor market frictions. As early as [Pissarides \(1985\)](#), the interaction between taxes and labor market frictions is studied in an equilibrium unemployment model. More recently [Lehmann et al. \(2006, 2011\)](#); [Golosov et al. \(2013\)](#); [Schaal and Taschereau-Dumuouchel \(2014\)](#); [Lehmann et al. \(2016\)](#); [Kroft et al. \(2017\)](#) are some of the works that have dealt with this issue. These works abstract from intensive margin choices, while introducing novel margins of response and/or heterogeneity. Whether a random search or directed search is considered is also an important difference across works.

In what we think is one of the most innovative contribution to the area, [Lehmann et al. \(2006\)](#), apply a mechanism design approach to an optimal taxation problem in a labor market characterized by random search. The assumption of ex-post wage bargain leads to surpluses for both the firm and the worker. The mechanism must extract information from the pair which, however cooperative, still have conflicting interests.

We, in contrast, make a minimum departure from [Mirrlees's](#) classic work to assess how labor *market frictions alone* affect optimal tax prescriptions and create extensive margin distortions. Moreover, we consider competitive search, which eliminates all surplus beyond the informational rents that accrue to workers.

Some of the aforementioned works also rely on a perturbation/sufficient statistic approach, making it hard to assess whether all relevant policy instruments are used. We, in contrast follow a mechanism design approach which allows us to assess what is feasible under the informational restrictions that real world governments face.

¹In some works, a fraction of workers does not work, e.g., [Jacquet et al. \(2013\)](#), but non-employment is voluntary.

1 Environment

The consumer/worker side of the economy is as in [Mirrlees \(1971\)](#). There is a continuum of agents with preferences defined over consumption, c , and effort, n , represented by

$$U(c, z, \theta) := \varphi(c) - \theta\eta(n),$$

where $\varphi : \mathbb{R}_+ \mapsto \mathbb{R}$ is a smooth, increasing, and strictly concave function and $\eta : \mathbb{R}_+ \mapsto \mathbb{R}$ is a smooth, increasing, and strictly convex function.

Heterogeneity across agents is captured by the disutility of effort parameter θ , distributed in the compact interval $[\underline{\theta}, \bar{\theta}]$ with $\underline{\theta} > 0$. We assume that θ is private information.

Our model differs from [Mirrlees'](#) in the treatment of the production side of the economy. While we retain the assumption that one unit of effort produces one unit of output, z , measured in units of consumption, we assume that opening a vacancy is costly for the firm. More precisely the firm must pay a cost $\kappa > 0$, also measured in units of consumption, for each job vacancy it creates.

A job opportunity is a contract specifying an output, z , to be produced and a payment y to which the worker is entitled upon producing z . We endow each firm with the ability to commit to a contract offer.² When deciding on which and how many work contracts to offer each firm takes as given, for each type of contract the ratio λ of workers applying for the same job and vacancies.

A worker who applies to a given job has a probability p of finding a job, where p is a function of λ . Intuitively, one can imagine that it is not too hard for a worker to observe a contract offer. But, provided that all other workers are able to observe the same offer, and in the absence of a centralized mechanism coordinating search efforts, nothing precludes multiple workers to reach for the same offers, leading some workers to remain unemployed.

In most of what follows, with some abuse in notation, we consider the inverse function $\lambda = \lambda(p)$. In this case, $\lambda(p)^{-1}$ maps the number of vacancies per worker which is necessary for the probability that a job applicant to find a job to be p . We assume that $\vartheta(p) := \lambda(p)^{-1}$ is a strictly convex, strictly increasing and twice differentiable function satisfying $\vartheta(0) = 0$, $\lim_{p \uparrow 1} \vartheta(p) = \infty$.

The economy is also inhabited by a benevolent planner/government who de-

²There is no renegotiation after a match is realized, which means that the wage compression effect highlighted in [Hungerbühler et al. \(2006\)](#), [Lehmann et al. \(2016\)](#) does not play a role here.

signs redistributive and insurance policies to maximize a social objective to be specified. We assume that the planner observes p , z and y , but not θ . In Section 3, when we take on the issue of implementation, we discuss the importance of observing these variables.

2 The Utilitarian Optimum

In this section we solve for the Utilitarian and Rawlsian optima. Our focus on these objectives is due to the fact that there are the objectives for which characterization results are better known.

Planner's Problem To characterize the constrained efficient allocation we consider a mechanism under which each agent announces his type, θ , and is assigned: a specific firm/market to which apply and how much to produce, $z(\theta)$, in case he receives a job offer. Associated with each firm/market is a probability of receiving a job offer, $p(\theta)$, a consumption level $c(\theta)$ conditional on working (and producing $z(\theta)$) and an unemployment benefit $c^u(\theta)$.

As it turns it is simpler to work with the transformations: $u(\theta) := \varphi(c(\theta))$, $h(\theta) := \eta(z(\theta))$. Define, in this case $C(\varphi(c)) := c$, $N(\eta(z)) := z$, and $\underline{u}(\theta) = \varphi(c^u(\theta))$.

Let us, for the moment consider the case in which job offers are observed by the planner, who has the instruments to force an agent who gets an offer to accept it. In this case, the planners problem is

$$\max \int [p(\theta) (u(\theta) - \theta h(\theta)) + (1 - p(\theta))\underline{u}(\theta)] f(\theta) d\theta$$

subject to

$$\int \left\{ p(\theta) [N(h(\theta)) - C(u(\theta))] - (1 - p(\theta))C(\underline{u}(\theta)) - \frac{\kappa}{\lambda(p(\theta))} \right\} f(\theta) d\theta,$$

$$w(\theta) := [p(\theta) (u(\theta) - \theta h(\theta)) + (1 - p(\theta))\underline{u}(\theta)],$$

$$\dot{w}(\theta) = -p(\theta)h(\theta),$$

and

$$p(\theta)h(\theta) \text{ decreasing.}$$

It is not hard to show that, the solution to this program leads to $c(\theta) = c^u(\theta)$

for all θ . This very unrealistic feature of optimal policy is due to the fact that we have assumed away the moral hazard aspect of unemployment insurance. We, therefore, take into account this aspect of the problem by assuming that the planner does not observe whether a job offer has been received by the agent.

If the planner does not observe whether a job offer has been received by the agent. In this case, the possibility of a double deviation arises. A type θ agent can claim to be a $\hat{\theta}$ type and reject all job offers he receives, thus guaranteeing himself a utility $\underline{u}(\hat{\theta})$. To avoid such double deviation, we add the constraint,

$$w(\theta) \geq \sup_{\tilde{\theta}} \underline{u}(\tilde{\theta}).$$

to the planner's program.

Claim 2.1. *If the planner does not observe whether job offers are received by the agents, then in any solution to the planner's problem we have $\underline{u}(\theta) = \underline{u}$ for almost every θ . Moreover, if we let $\hat{\theta}$ be the largest type such θ that have $h(\theta) > 0$ then*

$$\underline{u} = \left(u(\hat{\theta}) - \hat{\theta}h(\hat{\theta}) \right).$$

Since $w(\theta)$ is decreasing in θ , a necessary and sufficient condition for an allocation to be implementable is $w(\bar{\theta}) \geq \sup_{\bar{\theta}} \underline{u}(\bar{\theta})$.

Planner's Problem Dual Due to Claim 2.1, we know that the unemployment benefit must be independent of θ at the solution for the planner's. For any \underline{u} we can, therefore define the following dual program, \mathcal{P}^u ,

$$\max_{h(\theta), u(\theta), p(\theta)} \int \left\{ p(\theta) [N(h(\theta)) - C(u(\theta))] - (1 - p(\theta))C(\underline{u}) - \frac{\kappa}{\lambda(p(\theta))} \right\} f(\theta) d\theta,$$

subject to

$$\int w(\theta) f(\theta) d\theta \geq A, \tag{2.1}$$

$$\dot{w}(\theta) = -p(\theta)h(\theta), \tag{2.2}$$

$$w(\bar{\theta}) = \bar{u}, \tag{2.3}$$

and

$$p(\theta)h(\theta) \text{ decreasing.} \tag{2.4}$$

Notice that the indirect utility of a type θ worker is

$$w(\theta) = p(\theta) [u(\theta) - \theta h(\theta)] + (1 - p(\theta)) \underline{u},$$

which implies that

$$u(\theta) = \frac{w(\theta) - \underline{u}}{p(\theta)} + \theta h(\theta) + \underline{u}.$$

Hence, the planner's dual problem may be written as an optimal control program, with $h(\theta)$ and $p(\theta)$ as controls, $w(\theta)$, the state variable, and $\mu(\theta)$ the associated co-state.

We will restrict our attention to C^2 solutions which satisfy the monotonicity condition, and ignore the possibility of bunching.

It is then possible to show that the sign of the labor wedge, $N'(h(\theta)) - \theta C'(u(\theta))$, is pinned down by the sign of $\mu(\theta)$ through the first order condition with respect to $h(\theta)$,

$$[N'(h(\theta)) - \theta C'(u(\theta))] f(\theta) = \mu(\theta).$$

The same is true for the sign of the extensive margin wedge,

$$N(h(\theta)) - C(u(\theta)) + C(\underline{u}) + C'(u(\theta)) \frac{w(\theta) - \underline{u}}{p(\theta)} - \kappa \vartheta'(p(\theta)),$$

but now through the first order condition with respect to $p(\theta)$,

$$\left[N(h(\theta)) - C(u(\theta)) + C(\underline{u}) + C'(u(\theta)) \frac{w(\theta) - \underline{u}}{p(\theta)} - \kappa \vartheta'(p(\theta)) \right] f(\theta) = \mu(\theta) h(\theta),$$

To sign both wedges we must, therefore, assess the sign of $\mu(\theta)$. Unfortunately, this is not as simple as in the [Mirrlees'](#) case. Whereas in the [Mirrlees'](#) model implementation requires monotonicity in $h(\theta)$, hence in $u(\theta)$, here an implementable allocation must be monotonic in $p(\theta)h(\theta)$, not necessarily in each one separately. Indeed, the following is a very simple example of an economy which equilibrium is not monotonic in p : there exist θ' and θ , $\theta' > \theta$, such that $p(\theta') > p(\theta)$. As it turns, monotonicity in $u(\theta)$ is what we need to determine the sign of $\mu(\theta)$, so, following this example, Proposition 1, shows the Utilitarian optimum is monotonic in all variables.

Example 2.1. Assume that preferences are of the form

$$\ln c - \theta \frac{n^{1+\gamma}}{1+\gamma}$$

We assume that $p(\lambda) = (1 + \lambda)^{-1}$, which allow us to write the zero profit condition as $z - y = \kappa(1 - p)^{-1}$, where y denotes the agents earnings and z his output.

The tax schedule is piecewise linear do that budget sets are of the form

$$c = y(1 - \chi_{y \geq \bar{y}}\tau) + \chi_{y \geq \bar{y}}\tau\bar{y},$$

where $\chi_{y \geq \bar{y}}$ is an indicator function taking the value 1 if $y \geq \bar{y}$ and 0, otherwise. for

$$y = n - \frac{\kappa}{1 - p} \geq \bar{y}.$$

Hence, for a given p , the agent's intensive margin choice is the solution to the problem

$$\max_z \ln \left(\left(z - \frac{\kappa}{1 - p} \right) (1 - \chi_{y \geq \bar{y}}\tau) + \chi_{y \geq \bar{y}}\tau\bar{y} \right) - \theta \frac{z^{1+\gamma}}{1 + \gamma}$$

For $\gamma = 1$ the solution is

$$z(\theta) = \frac{1}{2} \left[-\frac{B}{1 - \chi_{y \geq \bar{y}}\tau} + \sqrt{\left(\frac{B}{1 - \chi_{y \geq \bar{y}}\tau} \right)^2 + \frac{4}{\theta}} \right]$$

for

$$B = \frac{-\kappa(1 - \chi_{y \geq \bar{y}}\tau)}{1 - p} + \chi_{y \geq \bar{y}}\tau\bar{y}.$$

Figure ?? displays the equilibrium functions $\bar{z}(\theta)$ and $\bar{p}(\theta)$ for this economy, under the parametrization

In the appendix we show that at the optimum $u(\theta)$ is decreasing in θ .

Proposition 1. At the Utilitarian optimum, $u(\theta)$, $h(\theta)$ and $p(\theta)$ are decreasing in θ .

Proposition 1 allows us to prove that $\mu(\theta) \geq 0$ for all θ . Note that incentive compatibility per se only implies that $p(\theta)h(\theta)$ are decreasing in θ , it is optimality for the Utilitarian objective which gives us monotonicity in each one separately.

Proposition 2. *The allocation $(p^*(\theta), h^*(\theta), u^*(\theta))_\theta$ that solves the planner's Utilitarian program is such that*

$$N'(h^*(\theta)) > \theta C'(u^*(\theta))$$

and

$$N(h^*(\theta)) - C(u^*(\theta)) + C(\underline{u}) + C'(u^*(\theta)) \frac{w^*(\theta) - \underline{u}}{p^*(\theta)} \geq \kappa \vartheta'(p^*(\theta)).$$

Proposition 2 guarantees that, at the optimum, $h(\theta)$ and $p(\theta)$ are locally distorted downward. That is, if it were not for incentive constraints it would be efficient to slightly increase both. The proposition says nothing about how $h(\theta)$ and $p(\theta)$ compare to first best allocations. We shall now show that, if preferences are iso-elastic, $p^*(\theta)$, $h^*(\theta)$, and $u^*(\theta)$ are all distorted downwards.

Optimal Distortions For each productivity type, θ , consider the program \mathcal{P}° below,

$$\min_{p, h, u} p[C(u) - N(h)] + (1 - p)C(\underline{u}) + \frac{\kappa}{\lambda(p)}$$

s.t.,

$$p[u - \theta h - \underline{u}] \geq w^*(\theta) - \underline{u},$$

where $w^*(\theta)$ is the utility attained by type- θ agents at the Utilitarian optimum. Let $(p^\circ(\theta), h^\circ(\theta), u^\circ(\theta))$ denote its solution.

The no distortion condition for the $\underline{\theta}$ -type leads to

$$(p^\circ(\underline{\theta}), h^\circ(\underline{\theta}), u^\circ(\underline{\theta})) = (p^*(\underline{\theta}), h^*(\underline{\theta}), u^*(\underline{\theta})).$$

The question we want to address next is whether we can say anything about how $(p^\circ(\theta), h^\circ(\theta), u^\circ(\theta))$ and $(p^*(\underline{\theta}), h^*(\underline{\theta}), u^*(\underline{\theta}))$ relate when $\theta \neq \underline{\theta}$.

Equation (C.1) and the first order conditions for the optimization problem above lead to

$$\frac{C'(u^*(\theta))}{N'(h^*(\theta))} < \frac{C'(u^\circ(\theta))}{N'(h^\circ(\theta))}. \quad (2.5)$$

Immediate from (2.5) are the facts that $h^*(\theta) > h^\circ(\theta) \Rightarrow u^*(\theta) < u^\circ(\theta)$ and $u^*(\theta) > u^\circ(\theta) \Rightarrow h^*(\theta) < h^\circ(\theta)$.

Noting that $p^*(\theta)[u^*(\theta) - \theta h^*(\theta) - \underline{u}] = p^\circ(\theta)[u^\circ(\theta) - \theta h^\circ(\theta) - \underline{u}]$, it is not hard to see that $u^*(\theta) > u^\circ(\theta) \Rightarrow p^*(\theta) < p^\circ(\theta)$ and $h^*(\theta) > h^\circ(\theta) \Rightarrow p^*(\theta) < p^\circ(\theta)$. Finally, note that $p^*(\theta) \geq p^\circ(\theta) \Rightarrow u^*(\theta) < u^\circ(\theta) \Rightarrow h^*(\theta) < h^\circ(\theta)$. If p is distorted upwards, then consumption must be distorted downwards.

Combining the first order conditions (C.1) and (C.2) we get

$$N(h^*(\theta)) - N'(h^*(\theta))h^*(\theta) - \left[C(u^*(\theta)) - C(\underline{u}) - C'(u^*(\theta))[u^*(\theta) - \underline{u}] \right] = \kappa \vartheta'(p^*(\theta)), \quad (2.6)$$

Analogously, by combining the first order conditions from program \mathcal{P}° , we obtain

$$N(h^\circ(\theta)) - N'(h^\circ(\theta))h^\circ(\theta) - \left[C(u^\circ(\theta)) - C(\underline{u}) - C'(u^\circ(\theta))[u^\circ(\theta) - \underline{u}] \right] = \kappa \vartheta'(p^\circ(\theta)). \quad (2.7)$$

The same expressions characterize the solution for the first and second best programs. Of course there are many different combinations of u , h and p that solve them. Indeed, for each θ , these variables are related through $p^*(\theta)[u^*(\theta) - \theta h^*(\theta) - \underline{u}]$ in the second best, and $p^\circ(\theta)[u^\circ(\theta) - \theta h^\circ(\theta) - \underline{u}]$ in the first best. Additional restrictions are imposed by optimality in both problems and incentive compatibility for the second best program.

Let us restrict ourselves to the case of iso-elastic preferences. In this case, $C'(u)u = A_c C(u)$ for some constant, A_c such that $0 < A_c < 1$, and $N'(h)h = A_n N(h)$ for some constant, A_n such that $A_n > 1$. Equation (2.6) becomes

$$N(h^*(\theta)) [1 - A_n] - \left[[C(u^*(\theta)) - C(\underline{u})] [1 - A_c] + [C'(u^*(\theta)) - C'(\underline{u})] \underline{u} \right] = \kappa \vartheta'(p^*(\theta)),$$

with an identical expression substituting for (2.7).

Now, assume that $p^*(\theta) \geq p^\circ(\theta)$. As we have seen this implies $u^*(\theta) < u^\circ(\theta)$ and $h^*(\theta) < h^\circ(\theta)$. But, in this case, either equality (2.6) or equality (2.7) is disrespected. In other words, at the optimum, $p^*(\theta) < p^\circ(\theta)$.

But, from $p^*(\theta)[u^*(\theta) - \theta h^*(\theta) - \underline{u}] = p^\circ(\theta)[u^\circ(\theta) - \theta h^\circ(\theta) - \underline{u}]$, if $p^*(\theta) < p^\circ(\theta)$, then $h^*(\theta) > h^\circ(\theta) \Rightarrow u^*(\theta) > u^\circ(\theta)$ which violates (2.5). Hence, $h^*(\theta) < h^\circ(\theta)$.

2.1 The Rawlsian case

In the Rawlsian case, the dual program, \mathcal{P}^r is identical to program \mathcal{P}^r , except that $w(\bar{\theta}) \geq \bar{w}$ substitutes for (2.1) for some 'exogenous' \bar{w} .

Proposition 3. *The allocation $(p^r(\theta), h^r(\theta), u^r(\theta))_\theta$ that solves the planner's Rawlsian program is such that*

$$N'(h^r(\theta)) \geq \theta C'(u^r(\theta))$$

and

$$N(h^r(\theta)) - C(u^r(\theta)) + C(\underline{u}) + C^r(u^r(\theta)) \frac{w^r(\theta) - \underline{u}}{p^r(\theta)} \geq \kappa \vartheta'(p^r(\theta)).$$

The proof is almost identical to the one in the Utilitarian case, so we omit for brevity.

3 Implementation via labor income taxes?

To assess whether optimal income taxation suffices we consider the firm's decentralized program.

Assume that only a labor income tax, $T(\cdot)$, is in place. We may implicitly define the function $\chi : \mathbb{U} \mapsto R$ through, $\varphi(\chi(u) - T(\chi(u))) = u$ for all $u \in \mathbb{U}$.

In this case, a firm which decides to offer a contract to a type θ solves a problem of the form

$$\min_{u, h} \lambda(p)p [\chi(u) - N(h)] + \kappa$$

s.t.

$$p [u - \theta h - \underline{u}] + \underline{u} \geq w(\theta). \quad (3.1)$$

The firm's optimization problem is, therefore, very similar to the autarky program. The difference is only that $\chi(\cdot)$ substitutes for $C(\cdot)$ as the cost function associated with the utility from consumption u . $\chi(\cdot)$ differs from $C(\cdot)$ because it takes into account the labor income tax schedule set in place by the government.

Also note that no incentive constraints are imposed on the firm. In fact, this is a private values environment in a competitive setting – see [Pouyet et al. \(2008\)](#). The firm's profit is independent of a worker's type, *conditional on a labor contract* (z, y) .

As it turns, to find the equilibrium allocation it is simpler to rely on the fact that any equilibrium allocation solves for every θ the following problem

$$\max_{p, h, u} p [u - \theta h - \underline{u}]$$

subject to

$$N(h) - \chi(u) \geq \frac{\kappa}{\lambda(p)p},$$

which we shall denote program \mathcal{P}^{EQ} .

We can use the zero profit condition above along with the firms' first order conditions to show that the planner cannot implement the constrained efficient allocation derived in Section 2.

Proposition 4. *The allocation $(p^*(\theta), h^*(\theta), u^*(\theta))_\theta$ which solves the planner's Utilitarian program cannot be implemented using only a non-linear labor income tax schedule.*

To understand the rationale for this finding recall that Claim 2.1 has established that, at the optimum, $u(\bar{\theta}) - \bar{\theta}h(\bar{\theta}) = \underline{u}$. The agent is in this case indifferent between any two values of p .

Since $p(\theta)$ is decreasing in θ , then we know that $p(\bar{\theta}) < 1$. The profit of a firm hiring $\bar{\theta}$ types is

$$\lambda(\bar{\theta})p(\bar{\theta}) [N(h(\bar{\theta})) - \chi(u(\bar{\theta}))] - \kappa \geq 0 \Rightarrow N(h(\bar{\theta})) > \chi(u(\bar{\theta}))$$

If the firm offers a slightly higher utility $u(\bar{\theta}) + \epsilon$, such that $N(h(\bar{\theta})) > \chi(u(\bar{\theta}) + \epsilon)$, it will make the agent strictly prefer to be employed and will allow the firm to raise $p(\bar{\theta})$ as much as it wants thus increasing profits.³

Observing h Suffices for Implementation Next, we show that if the government observes h , or, equivalently, the agent's output, then the second-best can always be implementable. Notice that when $h \in \{h(\theta)\}_{\theta \in \Theta}$ is observable we can assume that the firm reports h to the government, the government charges the firm the total payment $Z(\theta)$ implicitly defined by:

$$\lambda(p(\theta))p(\theta) [N(h(\theta)) - Z(\theta)] - \kappa = 0. \quad (3.2)$$

The government in turn makes a transfer equal to $C(u(\theta))$ to a worker who produced $N(h(\theta))$. We can thus assume that each vacancy is associated with a production level $N(h(\theta))$ and, by the zero-profit above, the number of vacancies per worker of a firm requiring effort $h(\theta)$ is $\lambda(p(\theta))^{-1}$. Clearly, no firm nor worker has any incentive to deviate. Next, using (3.2) one verifies that the government's

³Needs only to choose ϵ and $\hat{p} < 1$ in such a way that

$$N(h(\bar{\theta})) - \chi(u(\bar{\theta}) + \epsilon) > \frac{\lambda(p(\bar{\theta}))p(\bar{\theta})}{\lambda(\hat{p})\hat{p}} [N(h(\bar{\theta})) - \chi(u(\bar{\theta}))].$$

revenue is:

$$\begin{aligned} & \int \{p(\theta) [Z(h(\theta)) - C(u(\theta))] - (1 - p(\theta))C(\underline{u})\} f(\theta)d\theta \\ &= \int \left\{ p(\theta) [Z(h(\theta)) - C(u(\theta))] - (1 - p(\theta))C(\underline{u}) - \frac{\kappa}{\lambda(p(\theta))} \right\} f(\theta)d\theta, \end{aligned}$$

which is the value attained in the second best.

4 Assessing the value of controlling p

Our goal in this section is to evaluate the potential gains from adopting an efficient policy. Toward this end we start with a status quo labor income tax schedule and let $(\omega(\theta))_\theta$ denote the utility attained by agents under such schedule.

Then we consider the following program,

$$\max_{h(\theta), u(\theta), p(\theta)} \int \left\{ p(\theta) [N(h(\theta)) - C(u(\theta))] - (1 - p(\theta))C(\underline{u}) - \frac{\kappa}{\lambda(p(\theta))} \right\} f(\theta)d\theta,$$

subject to

$$w(\theta) \geq \omega(\theta), \forall \theta, \quad (4.1)$$

$$\dot{w}(\theta) = -p(\theta)h(\theta), \forall \theta, \quad (4.2)$$

and

$$p(\theta)h(\theta) \text{ decreasing}. \quad (4.3)$$

Finally we compare the revenue raised under the planner's program with the revenue raised under the status quo schedule.

Ignoring bunching, the first order conditions are (C.1), (C.2), in the appendix, and

$$-\dot{\mu}(\theta) = -C'(u(\theta))f(\theta) + \tilde{\psi}(\theta),$$

where $\tilde{\psi}(\theta)$ is the multiplier associated with constraint (4.1).

Note that since

$$\theta C'(u(\theta)) = N'(h(\theta)),$$

we can express $u(\theta) = C'^{-1}(N'(h(\theta))/\theta)$.

At the optimum, constraint (4.1) binds and we have $-\dot{w}(\theta) = p(\theta)h(\theta)$. We may

in this case write

$$C'^{-1}(N'(h(\theta))/\theta) = -\frac{\omega(\theta) - \underline{u}}{\dot{\omega}(\theta)}h(\theta).$$

For any \underline{u} and a given path for $\omega(\theta)$, equation (4.4) is a function of $h(\theta)$ only. In other words, for a given \bar{u} we use (4.4) to solve for $h(\theta)$.

All we need is to determine \underline{u} . But note that at the optimum $u(\bar{\theta}) - \bar{\theta}h(\bar{\theta}) = \underline{u}$. Hence,

$$\omega(\bar{\theta}) = w(\bar{\theta}) = p(\bar{\theta})[u(\bar{\theta}) - \bar{\theta}h(\bar{\theta})] + [1 - p(\bar{\theta})]\underline{u} = \underline{u}.$$

That is, \underline{u} is pinned down at the optimum by the least productive agent's equilibrium utility, $\omega(\bar{\theta})$.

Combining (C.1) and (C.2), we have

$$\begin{aligned} N(h(\theta)) - C(u(\theta)) + C(\underline{u}) + C'(u(\theta)) \left[\frac{w(\theta) - \underline{u}}{p(\theta)} \right] \\ - \kappa \vartheta'(p(\theta)) = N'(h(\theta))h(\theta) - C'(u(\theta))\theta h(\theta) \quad (4.4) \end{aligned}$$

Since we have $h(\theta)$ we immediately obtain $p(\theta)$ and $u(\theta)$ using (4.2) and the equation above, respectively.

Hence, all we need to implement the assessment above is to find, for all θ , the equilibrium utility, $\omega(\theta)$.

4.1 Quantitative Assessment

Whether we can and under what assumptions we can recover the utility profile ($\omega(\theta)$) naturally depends on what we observe in the data.

We shall first consider the case in which only the distribution of earnings is observed. This will oblige us to make many parametric assumptions and solve for the equilibrium.

We parametrize the problem as follows. We assume $\varphi(c) = c^{1-\sigma}/(1-\sigma)$, $\eta(n) = n^{1+\gamma}/(1+\gamma)$ and $p = (1+\lambda)^{-1}$, which implies $\lambda(p) = 1/p - 1$. Moreover, to parametrize the US tax system, we rely on the functional form $T(y) = y - \xi y^{1-\tau}$ proposed by [Musgrave \(1959\)](#); [Feldstein \(1969\)](#) – and used by [Bénabou \(2000, 2002\)](#); [Heathcote et al. \(2017\)](#), to name a few.

As we have seen in Section 3, page 11, this characterization is easier if we rely on program \mathcal{P}^{EQ} . That is, an equilibrium allocation solves, for a firm participating

in the θ -market the dual program,

$$\max_{p,h,u} p [u - \theta h - \underline{u}]$$

subject to

$$N(h) - \chi(u) \geq \frac{\kappa}{\lambda(p)p}.$$

Solving this program for the functional forms we are using here, we can find an analytic expression for θ as a function of \underline{u} , κ , and y .

It is a maintained assumption of this work that earnings are observed. As for κ and \underline{u} , if in the data low productivity agents are non-participants, then, we can use the lowest skill for which agents participate, say θ_0 , to pin down \underline{u} through $u(\theta_0) - \theta_0 h(\theta_0) = \underline{u}$. As for κ note that, for each θ ,

$$\left[(1 - \tau) \zeta^{1-\sigma} y^{(1-\tau)(1-\sigma)-1} \right]^{\frac{1}{\gamma}} \theta^{-\frac{1}{\gamma}} - y = \frac{\kappa}{1-p}.$$

If we do not observe $p(\theta)$ for each θ , but only on average, we can pick κ to match this average participation rate. If, however, all productivity types participate, then κ and \underline{u} must be simultaneously determined by combining these two expressions.

4.2 Numeric results

TO BE DONE

5 Conclusion

TO BE DONE

References

Bénabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review* 90(1), 96–129.

Bénabou, R. (2002). Tax and education policy in a heterogeneous-agent economy: What levels of redistribution maximize growth and efficiency? *Econometrica* 70(2), 481 – 517.

- Feldstein, M. (1969). The effects of taxation on risk taking. *Journal of Political Economy* 77(5), 755 – 764.
- Golosov, M., P. Maziero, and G. Menzio (2013). Taxation and redistribution of residual income inequality,. *Journal of Political Economy*, 121(6), pp.116-1204. 121(6), 1160–1204.
- Heathcote, J., K. Storesletten, and G. Violante (2017). Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics* forthcoming.
- Hungerbühler, M., E. Lehmann, A. Parmentier, and B. Van der Linden (2006). Optimal redistributive taxation in an equilibrium search model. *Review of Economic Studies* 73, 743 – 767.
- Jacquet, L., E. Lehmann, and B. Van der Linden (2013). The optimal marginal tax rates with both extensive and intensive responses. *Journal of Economic Theory* 148(5), 1770 – 1805.
- Kroft, K., K. Kucko, E. Lehmann, and J. Schmeider (2017, September). Optimal income taxation with unemployment and wage responses: A sufficient statistics approach. mimeo.
- Lehmann, E., M. Hungerbühler, A. Parmentier, and B. Van der Linden (2006). Optimal redistributive taxation in a search equilibrium model. *Review of Economics Studies* 73(3), 743–768.
- Lehmann, E., B. Lucifora, S. Moriconi, and B. Van der Linden (2016). Beyond the labour income tax wedge: The unemployment-reducing effect of tax progressivity. *International Tax and Public Finance* 23(3), 454–489.
- Lehmann, E., A. Parmentier, and B. Van der Linden (2011). Optimal income taxation with endogenous participation and search unemployment. *Journal of Public Economic* 95(1523 – 1537).
- Mirrlees, J. A. (1971). An exploration in the theory of optimal income taxation. *Review of Economic Studies* 38, 175–208.
- Musgrave, R. A. (1959). *The Theory of Public Finance*. New York.
- Pissarides, C. A. (1985). Taxes, subsidies and equilibrium unemployment. *Review of Economic Studies* 52(1), 121 – 133.
- Pouyet, J., B. Salanié, and F. Salanié (2008). On competitive equilibria with asymmetric information. *The B.E. Journal of Theoretical Economics* 8(1 (Topics)), Article 13.
- Schaal, E. and M. Taschereau-Dumuouchel (2014). Optimal redistributive policy in a labor market with search and endogenous participation. mimeo.

A Proofs

Proof of Claim 2.1. Let $\underline{u} := \sup \underline{u}(\tilde{\theta})$ and suppose towards a contradiction that we can find $\tilde{\theta}$ such that

$$\underline{u}(\tilde{\theta}) < \underline{u}.$$

We have

$$[p(\theta)(u(\theta) - \theta h(\theta)) + (1 - p(\theta))\underline{u}(\theta)] \geq \underline{u},$$

which implies that $\underline{u}(\theta) < \underline{u} < u(\theta)$. Hence since the function C is convex one can construct a least costly allocation by decreasing $u(\theta)$ by $(1 - p(\theta))\varepsilon$ and increasing $\underline{u}(\theta)$ by $p(\theta)\varepsilon$. If this change is possible in a positive-measure set of types $\tilde{\Theta} \subset \Theta$ then it saves a positive amount of resources, which can then be redistributed by increasing the resulting utilities $u^*(\theta)$ and $\underline{u}^*(\theta)$ uniformly by some $\chi > 0$.

Next assume that $\underline{u} < (u(\hat{\theta}) - \hat{\theta}h(\hat{\theta}))$ and notice that this implies that $\hat{\theta} = \bar{\theta}$ and hence since $w(\theta)$ is decreasing, we can find $\varepsilon > 0$ such that $\underline{u} + \varepsilon < u(\theta)$ and $w(\theta) > \underline{u} + \varepsilon$ for every θ . But then, for every θ one can find $\eta > 0$ such that the allocation can be made less costly by decreasing $u(\theta)$ by $(1 - p(\theta))\eta$ and increasing $\underline{u}(\theta)$ by $p(\theta)\eta$, a contradiction. \square

Proof of Proposition 4. Differentiating the zero profit condition,

$$N(h(\theta)) - \chi(u(\theta)) = \frac{\kappa}{\lambda(p(\theta))p(\theta)},$$

we get

$$N'(h(\theta))\dot{h}(\theta) - \chi'(u(\theta))\dot{u}(\theta) = -\frac{\kappa[\lambda'(p(\theta))p(\theta) + \lambda(p(\theta))]\dot{p}(\theta)}{[\lambda(p(\theta))p(\theta)]^2}.$$

Next, using the firms' first order conditions,

$$N'(h(\theta)) \left[\dot{h}(\theta) - \frac{\dot{u}(\theta)}{\theta} \right] = -\frac{\kappa[\lambda'(p(\theta))p(\theta) + \lambda(p(\theta))]\dot{p}(\theta)}{[\lambda(p(\theta))p(\theta)]^2}.$$

Finally, using the agents' envelope and the definition of $w(\theta)$,

$$N'(h(\theta)) \left[\frac{\dot{p}(\theta)}{\theta p(\theta)} \left[u(\theta) - \theta h(\theta) - \underline{u} \right] \right] = -\frac{\kappa[\lambda'(p(\theta))p(\theta) + \lambda(p(\theta))]\dot{p}(\theta)}{[\lambda(p(\theta))p(\theta)]^2}.$$

This must hold for all agents.

But, we know from Claim 2.1 that there is at least one type $\hat{\theta}$ such that

$$\underline{u} = \left(u(\hat{\theta}) - \hat{\theta}h(\hat{\theta}) \right),$$

which implies

$$\frac{\kappa \left[\lambda'(p(\hat{\theta}))p(\hat{\theta}) + \lambda(p(\hat{\theta})) \right] \dot{p}(\hat{\theta})}{[\lambda(p(\hat{\theta}))p(\hat{\theta})]^2} = 0.$$

This condition cannot be satisfied for any $p(\theta) < 1$.

□

B Lemmata

Lemma 1. *At the optimum $p(\theta)$ is decreasing in θ .*

Proof of Lemma 1. The agent's first order condition,

$$\dot{u}(\theta) - \theta \dot{h}(\theta) = -\frac{\dot{p}(\theta)}{p(\theta)} [u(\theta) - \theta h(\theta) - \underline{u}] \quad (\text{B.1})$$

combined with the agent's second order condition,

$$\dot{p}(\theta)h(\theta) + p(\theta)\dot{h}(\theta) \leq 0 \Rightarrow \frac{\dot{p}(\theta)}{p(\theta)} \leq -\frac{\dot{h}(\theta)}{h(\theta)},$$

implies

$$\dot{u}(\theta) - \theta \dot{h}(\theta) \frac{\dot{h}(\theta)}{h(\theta)} = -\frac{\dot{p}(\theta)}{p(\theta)} [u(\theta) - \underline{u}] + \frac{\dot{p}(\theta)}{p(\theta)} \theta \dot{h}(\theta),$$

or

$$\dot{u}(\theta) \leq -\frac{\dot{p}(\theta)}{p(\theta)} [u(\theta) - \underline{u}]. \quad (\text{B.2})$$

Recall that the first order condition with respect to $p(\theta)$ is

$$\left\{ N(h(\theta)) - C(u(\theta)) + C(\underline{u}) + C'(u(\theta)) [u(\theta) - \theta h(\theta) - \underline{u}] \right. \\ \left. - \kappa \vartheta'(p(\theta)) \right\} f(\theta) = \mu(\theta)h(\theta) \quad (\text{B.3})$$

The first order condition with respect to $h(\theta)$ is

$$[N'(h(\theta)) - \theta C'(u(\theta))] f(\theta) = \mu(\theta),$$

which, after multiplying by $h(\theta)$ becomes

$$[N'(h(\theta))h(\theta) - \theta h(\theta)C'(u(\theta))] f(\theta) = \mu(\theta)h(\theta). \quad (\text{B.4})$$

Substituting (B.4) in (B.3), we get

$$\begin{aligned} N(h(\theta)) - C(u(\theta)) + C(\underline{u}) + C'(u(\theta)) [u(\theta) - \theta h(\theta) - \underline{u}] \\ - \kappa \vartheta'(p(\theta)) = N'(h(\theta))h(\theta) - \theta h(\theta)C'(u(\theta)), \end{aligned}$$

which simplifies to

$$\begin{aligned} N(h(\theta)) - N'(h(\theta))h(\theta) - C(u(\theta)) + C(\underline{u}) + \\ C'(u(\theta)) [u(\theta) - \underline{u}] = \kappa \vartheta'(p(\theta)). \quad (\text{B.5}) \end{aligned}$$

Differentiating (B.5) with respect to θ yields

$$\begin{aligned} N'(h(\theta))\dot{h}(\theta) - N'(h(\theta))\dot{h}(\theta) - N''(h(\theta))\dot{h}(\theta)h(\theta) - C'(u(\theta))\dot{u}(\theta) + \\ C'(u(\theta))\dot{u}(\theta) + C''(u(\theta)) [u(\theta) - \underline{u}] \dot{u}(\theta) = \kappa \vartheta''(p(\theta)), \end{aligned}$$

which simplifies to

$$-N''(h(\theta))h(\theta)\dot{h}(\theta) + C''(u(\theta)) [u(\theta) - \underline{u}] \dot{u}(\theta) = \kappa \vartheta''(p(\theta))\dot{p}(\theta).$$

Hence assume towards a contradiction that we find an interior θ such that $\dot{p}(\theta) > 0$. Since

$$\frac{\dot{p}(\theta)}{p(\theta)} \leq -\frac{\dot{h}(\theta)}{h(\theta)}$$

we have

$$\frac{\dot{h}(\theta)}{h(\theta)} \leq -\frac{\dot{p}(\theta)}{p(\theta)} < 0.$$

Therefore,

$$\dot{u}(\theta) = \frac{\kappa \vartheta''(p(\theta))\dot{p}(\theta) + N''(h(\theta))h(\theta)\dot{h}(\theta)}{C''(u(\theta)) [u(\theta) - \underline{u}]} > 0. \quad (\text{B.6})$$

But recall that (B.2) implies

$$\dot{u}(\theta) \leq -\frac{\dot{p}(\theta)}{p(\theta)} [u(\theta) - \underline{u}] < 0.$$

A contradiction, which completes the proof. \square

Lemma 2. *At the optimum, $\mu(\theta) > 0$ for every $\theta > \underline{\theta}$.*

Proof. Assume that we can find a maximal interval (θ^*, θ^{**}) such that $\mu(\theta) < 0$ for every θ in this interval. From the maximality of this interval, we have: $\mu(\theta^*) = \mu(\theta^{**}) = 0$ and $\dot{\mu}(\theta^*) \leq 0$ and $\dot{\mu}(\theta^{**}) \geq 0$. Hence, since

$$\dot{\mu}(\theta) = -[\psi - C'(u(\theta))] f(\theta),$$

we have

$$C'(u(\theta^*)) \leq C'(u(\theta^{**})) \Rightarrow u(\theta^*) \leq u(\theta^{**}),$$

from the strict convexity of C . We will show that this condition is impossible.

We are now ready to complete the proof, contradicting the assumption that $u(\theta^*) \leq u(\theta^{**})$, $\mu(\theta^*) = \mu(\theta^{**})$ and $\mu(\theta) < 0$ for every $\theta \in (\theta^*, \theta^{**})$. First, notice that we have:

$$N'(h(\theta^*)) - \theta^* C'(u(\theta^*)) = N'(h(\theta^{**})) - \theta^{**} C'(u(\theta^{**})) = 0 \quad (\text{B.7})$$

and $N'(h(\theta)) - \theta C'(u(\theta)) < 0$ for every $\theta \in (\theta^*, \theta^{**})$. Define $T(\theta) := N(h(\theta)) - C(u(\theta))$ and notice that (B.7) implies that

$$T(\theta^*) = T(\theta^{**}) = 0. \quad (\text{B.8})$$

On the other hand, we have:

$$T(\theta^{**}) - T(\theta^*) = \int_{\theta^*}^{\theta^{**}} T'(\theta) d\theta = \int_{\theta^*}^{\theta^{**}} [N'(h(\theta))\dot{h}(\theta) - C'(u(\theta))\dot{u}(\theta)] d\theta.$$

Take any $\theta \in (\theta^*, \theta^{**})$. First, suppose $\dot{h}(\theta) \geq 0$. In this case, using

$$\dot{u}(\theta) = \theta \dot{h}(\theta) - \frac{\dot{p}(\theta)}{p(\theta)} [u(\theta) - \theta h(\theta) - \underline{u}]$$

we have

$$\begin{aligned} \left[N'(h(\theta))\dot{h}(\theta) - C'(u(\theta))\dot{u}(\theta) \right] &= [N'(h(\theta)) - C'(u(\theta))\theta] \dot{h}(\theta) + \\ &C'(u(\theta)) \frac{\dot{p}(\theta)}{p(\theta)} [u(\theta) - \theta h(\theta) - \underline{u}] < 0, \end{aligned}$$

because $[N'(h(\theta)) - C'(u(\theta))\theta] = T'(\theta) < 0$, $\dot{p}(\theta) < 0$ and $[u(\theta) - \theta h(\theta) - \underline{u}] > 0$.

Next, suppose $\dot{h}(\theta) < 0$. In this case, if $\dot{u}(\theta) \leq 0$, then

$$-C'(u(\theta))\dot{u}(\theta) \leq -N'(h(\theta)) \left(\frac{\dot{u}(\theta)}{\theta} \right).$$

Hence,

$$\begin{aligned} \left[N'(h(\theta))\dot{h}(\theta) - C'(u(\theta))\dot{u}(\theta) \right] &\leq \left[N'(h(\theta))\dot{h}(\theta) - N'(h(\theta)) \left(\frac{\dot{u}(\theta)}{\theta} \right) \right] \\ &= N'(h(\theta))\dot{h}(\theta) - N'(h(\theta)) \left(\frac{\theta\dot{h}(\theta) - \frac{\dot{p}(\theta)}{p(\theta)} [u(\theta) - \theta h(\theta) - \underline{u}]}{\theta} \right) \\ &= N'(h(\theta)) \left(\frac{\dot{p}(\theta)}{p(\theta)\theta} \right) [u(\theta) - \theta h(\theta) - \underline{u}] < 0. \end{aligned}$$

Finally, suppose that $\dot{h}(\theta) < 0$ and $\dot{u}(\theta) > 0$. In this case, we immediately have:

$$N'(h(\theta))\dot{h}(\theta) - C'(u(\theta))\dot{u}(\theta) < 0.$$

Hence we always have $T'(\theta) < 0$, which implies

$$\int_{\theta^*}^{\theta^{**}} T'(\theta) d\theta < 0$$

and contradicts (B.8). □

C Long Derivations

C.1 Deriving Expressions for the Wedges

The planner's dual problem may be written as the following optimal control program,

$$\max \int_{\underline{\theta}}^{\bar{\theta}} \left\{ p(\theta) \left[N(h(\theta)) - C \left(\frac{w(\theta) - \underline{u}}{p(\theta)} + \theta h(\theta) + \underline{u} \right) \right] - (1 - p(\theta))C(\underline{u}) - \frac{\kappa}{\lambda(p(\theta))} \right\} f(\theta) d\theta,$$

subject to

$$\begin{aligned} \int_{\underline{\theta}}^{\bar{\theta}} w(\theta) f(\theta) d\theta &\geq A, \\ \dot{w}(\theta) &= -p(\theta)h(\theta), \\ w(\bar{\theta}) &= \bar{u}, \end{aligned}$$

and

$$p(\theta)h(\theta) \text{ decreasing.}$$

Here, $h(\theta)$ and $p(\theta)$ are the controls and $w(\theta)$ is the state variable.

We will restrict our attention to C^2 solutions which satisfy the monotonicity condition. We can thus write the Lagrangean:

$$\begin{aligned} V(A, \underline{u}) = \max \int_{\underline{\theta}}^{\bar{\theta}} \left\{ p(\theta) \left[N(h(\theta)) - C \left(\frac{w(\theta) - \underline{u}}{p(\theta)} + \theta h(\theta) + \underline{u} \right) \right] - (1 - p(\theta))C(\underline{u}) - \kappa \vartheta(p(\theta)) + \psi [w(\theta) - A] \right\} f(\theta) d\theta - \mu(\theta)p(\theta)h(\theta). \end{aligned}$$

Ignoring bunching, the first order conditions are

$$p(\theta) [N'(h(\theta)) - \theta C'(u(\theta))] f(\theta) = \mu(\theta)p(\theta), \quad (\text{C.1})$$

$$\begin{aligned} \left[N(h(\theta)) - C(u(\theta)) + C(\underline{u}) + C'(u(\theta)) \frac{w(\theta) - \underline{u}}{p(\theta)} - \kappa \vartheta'(p(\theta)) \right] f(\theta) = \mu(\theta)h(\theta), \quad (\text{C.2}) \end{aligned}$$

and

$$-\dot{\mu}(\theta) = -C'(u(\theta))f(\theta) + \psi f(\theta). \quad (\text{C.3})$$

To sign the labor wedge in (C.1) we must assess the sign of $\mu(\theta)$. Integrating $-\dot{\mu}(\theta) = -C'(u(\theta))f(\theta) + \psi f(\theta)$ from $\underline{\theta}$ to $\bar{\theta}$, we obtain

$$-\int_{\underline{\theta}}^{\bar{\theta}} \dot{\mu}(\theta) d\theta = \int_{\underline{\theta}}^{\bar{\theta}} [\psi f(\theta) - C'(u(\theta))f(\theta)] d\theta,$$

or

$$\mu(\bar{\theta}) = -\int_{\underline{\theta}}^{\bar{\theta}} [\psi f(\theta) - C'(u(\theta))f(\theta)] d\theta + \mu(\underline{\theta}).$$

Since $\mu(\bar{\theta}) = 0$, we get

$$\mu(\bar{\theta}) = -\int_{\underline{\theta}}^{\bar{\theta}} [\psi f(\theta) - C'(u(\theta))f(\theta)] d\theta.$$

Recall that, since $w(\bar{\theta}) = \underline{u}$, we do not necessarily have $\mu(\bar{\theta}) = 0$. We would like to know the sign of $\mu(\bar{\theta})$. Notice that the allocation $(u^*(\theta) + x, \underline{u} + x, p(\theta), h(\theta))$ is always feasible for $|x| < \varepsilon$ for some $\varepsilon > 0$.

Therefore, we obtain

$$-\frac{\partial V(A, \underline{u})}{\partial A} = \psi = \int_{\underline{\theta}}^{\bar{\theta}} [p(\theta)C'(u(\theta)) + (1 - p(\theta))C'(\underline{u})] f(\theta) d\theta,$$

which finally implies

$$\begin{aligned} \mu(\bar{\theta}) &= -\int_{\underline{\theta}}^{\bar{\theta}} [p(\theta)C'(u(\theta)) + (1 - p(\theta))C'(\underline{u})] f(\theta) d\theta + \int_{\underline{\theta}}^{\bar{\theta}} C'(u(\theta))f(\theta) d\theta \\ &= \int_{\underline{\theta}}^{\bar{\theta}} (1 - p(\theta)) (C'(u(\theta)) - C'(\underline{u})) f(\theta) d\theta > 0. \end{aligned}$$

The least productive agent, faces a positive marginal tax rate. What can we say about the other agents? Integrating (C.3) now from $\underline{\theta}$ to θ yields

$$\mu(\theta) = -\int_{\underline{\theta}}^{\theta} [\psi f(a) - C'(u(a))f(a)] da + \mu(\underline{\theta}).$$

If $u(\theta)$ is increasing in θ , then we obtain $\mu(\theta) > 0$. Proposition 1 guarantees that

this is the case.

C.2 Deriving expression (??)

Using our preferred parametrization (and returning to the primal variables), it is then the case that an equilibrium allocation solves

$$\max_{p,z,y} p \left[\frac{1}{1-\sigma} (\zeta y^{1-\tau})^{1-\sigma} - \frac{\theta}{1+\gamma} z^{1+\gamma} - \underline{u} \right]$$

subject to

$$z - y \geq \frac{\kappa}{1-p}. \quad (\text{C.4})$$

The first order condition for this problem are

$$p(1-\tau)\zeta^{1-\sigma}y^{(1-\tau)(1-\sigma)-1} = \alpha(\theta), \quad (\text{C.5})$$

with respect to y ,

$$\theta p z^\gamma = \alpha(\theta), \quad (\text{C.6})$$

with respect to z , and

$$\left[\frac{1}{1-\sigma} (\zeta y^{1-\tau})^{1-\sigma} - \frac{\theta}{1+\gamma} z^{1+\gamma} - \underline{u} \right] = \alpha(\theta) \frac{\kappa}{(1-p)^2} \quad (\text{C.7})$$

with respect to p .

Using (C.5) and (C.6) we get

$$\left[(1-\tau)\zeta^{1-\sigma}y^{(1-\tau)(1-\sigma)-1} \right]^{\frac{1}{\gamma}} \theta^{-\frac{1}{\gamma}} = z \quad (\text{C.8})$$

which shows that no extra distortion, beyond that caused by labor income taxes, is introduced by the firm.

Multiplying (C.5) by $y/(1-\sigma)$, (C.6) by $z/(1+\gamma)$, and adding the two we get

$$p \left[\frac{(\zeta y^{1-\tau})^{1-\sigma}}{1-\sigma} - \theta \frac{z^{1+\gamma}}{1+\gamma} - \bar{u} \right] = \frac{\alpha(\theta)y}{1-\sigma} - \frac{\alpha(\theta)z}{1+\gamma} - p\bar{u},$$

Multiplying (C.5) by $y/(1-\tau)(1-\sigma)$, (C.6) by $z/(1+\gamma)$, and adding the two

we get

$$p \left[\frac{(\zeta y^{1-\tau})^{1-\sigma}}{1-\sigma} - \theta \frac{z^{1+\gamma}}{1+\gamma} - \bar{u} \right] = \frac{\alpha(\theta)y}{(1-\tau)(1-\sigma)} - \frac{\alpha(\theta)z}{1+\gamma} - p\bar{u},$$

which, using (C.7) can be written as

$$-\alpha(\theta) \frac{p\kappa}{(1-p)^2} = \frac{\alpha(\theta)y}{1-\sigma} - \frac{\alpha(\theta)z}{1+\gamma} - p\bar{u}.$$

$$\alpha(\theta) \frac{p\kappa}{(1-p)^2} = \frac{\alpha(\theta)y}{(1-\tau)(1-\sigma)} - \frac{\alpha(\theta)z}{1+\gamma} - p\bar{u}.$$

We can eliminate $\alpha(\theta)$ using (C.6):

$$\bar{u} = \theta z^\gamma \left[\frac{y}{1-\sigma} - \frac{z}{1+\gamma} + \frac{\kappa}{(1-p)^2} \right].$$

$$\bar{u} = \theta z^\gamma \left[\frac{y}{(1-\tau)(1-\sigma)} - \frac{z}{1+\gamma} + \frac{p\kappa}{(1-p)^2} \right].$$

Next, using the fact that constraint (C.4) is active at the optimum, we have

$$(z-y)^2 = \frac{\kappa^2}{(1-p)^2}$$

$$p = 1 - \frac{\kappa}{z-y}$$

which can be used to obtain wrong

$$\bar{u} = \theta z^\gamma \left[\frac{y}{1-\sigma} - \frac{z}{1+\gamma} + \frac{z^2 - 2zy + y^2}{\kappa} \right].$$

correct

$$\bar{u} = \theta z^\gamma \left[\frac{y}{(1-\tau)(1-\sigma)} - \frac{z}{1+\gamma} - \frac{(z-y)^2}{\kappa} + (z-y) \right].$$

or

$$\bar{u} = \theta z^\gamma \left[\frac{1-(1-\tau)(1-\sigma)}{(1-\tau)(1-\sigma)} y + \frac{\gamma}{1+\gamma} z - \frac{z^2}{\kappa} + \frac{2zy}{\kappa} - \frac{y^2}{\kappa} \right].$$

wrong

Using (C.8) we get

$$\bar{u} = \zeta^{1-\sigma} y^{(1-\tau)(1-\sigma)-1} \left[\frac{y}{1-\sigma} - \frac{\zeta^{\frac{1-\sigma}{\gamma}} y^{\frac{(1-\tau)(1-\sigma)-1}{\gamma}} \theta^{-\frac{1}{\gamma}}}{1+\gamma} + \frac{\zeta^{\frac{2(1-\sigma)}{\gamma}} y^{\frac{2(1-\tau)(1-\sigma)-1}{\gamma}} \theta^{-\frac{2}{\gamma}} - 2\zeta^{\frac{1-\sigma}{\gamma}} y^{\frac{(1-\tau)(1-\sigma)+\gamma-1}{\gamma}} \theta^{-\frac{1}{\gamma}} + y^2}{\kappa} \right] \quad (\text{C.9})$$

Let $A(y) := \zeta^{1-\sigma} y^{(1-\tau)(1-\sigma)-1}$, then we can write

$$\frac{\bar{u}}{A(y)} - \frac{y}{1-\sigma} - \frac{y^2}{\kappa} = \left[\frac{A(y)^{\frac{1}{\gamma}}}{1+\gamma} + \frac{2A(y)^{\frac{1}{\gamma}} y}{\kappa} \right] \theta^{-\frac{1}{\gamma}} + \frac{A(y)^{\frac{2}{\gamma}}}{\kappa} \theta^{-\frac{2}{\gamma}}.$$

correct

Let $A(y) := (1-\tau)\zeta^{1-\sigma} y^{(1-\tau)(1-\sigma)-1}$, then we can write

$$-\frac{\bar{u}}{A(y)} + \frac{1-(1-\tau)(1-\sigma)}{(1-\tau)(1-\sigma)} y - \frac{y^2}{\kappa} = - \left[\frac{\gamma}{1+\gamma} A(y)^{\frac{1}{\gamma}} + \frac{2A(y)^{\frac{1}{\gamma}} y}{\kappa} \right] \theta^{-\frac{1}{\gamma}} + \frac{A(y)^{\frac{2}{\gamma}}}{\kappa} \theta^{-\frac{2}{\gamma}}.$$

This is a second degree polynomial equation in $Z = \theta^{-\frac{1}{\gamma}}$ which has closed form solution.