Simple Labor Income Tax Systems with Endogenous Employment Contracts

Anqi Li[*] Yiqing Xing[†]

December, 2013

Abstract

In this paper, we use an anonymous and time-invariant progressive labor income tax system to implement the constrained social optimum in a setting where workers privately experience both persistent ability shocks and transient productivity shocks. We propose a framework to capture the interplay between welfare policies and firms' contractual choices. In particular, we use the tax system to achieve two goals: directly, to redistribute the life-cycle income between workers of different ability types; indirectly, to induce firms to absorb transient productivity shocks through efficiency wage contracts.

[*]Department of Economics, Washington University in St. Louis. angellianqi@gmail.com.

[†]Department of Economics, Stanford University. xingyq@stanford.edu. We have benefited from helpful conversations with Matt Jackson, Florian Scheuer and Ilya Segal. This work is very preliminary. Comments are highly welcomed.

1 Introduction

A successful welfare policy should implement efficient outcomes with simple policy instruments. At the first sight, these criteria seem to be conflicting because the environment we face are quite complicated. For one thing, we experience varying types of productivity shocks throughout our life cycles, some of which are frequent but have temporary effects (e.g., traffic, unanticipated difficulty, unexpected delay in inputs) whereas others are infrequent but carry long-lasting impacts (e.g., innate ability, handicap, sectoral unemployment). For another, these productivity shocks are often our private information. Together, these considerations suggest that the implementation of any interesting outcome would require complex policies such as personalized income tax systems that vary the tax rate from person to person and history by history. Nevertheless, real-world welfare policies are much simpler, as they are typically anonymous and considerably more stable than the changing environment in our daily lives. Does such simplicity come at the cost of efficiency? Put it differently, can we implement the constrained social optimum with simple policies? If so, under what conditions?

This paper provides affirmative answers to these questions in a setting where workers privately experience a persistent ability shock a la Mirrlees (1971) and many transient productivity shocks throughout their life cycles. The main result shows that under certain regularity conditions, we can implement the constrained social optimum with an anonymous and time-invariant progressive labor income tax system. Our key contribution is to let firms play an active role in the implementation strategy. Specifically, we use the labor income tax system to achieve two goals: directly, to redistribute the life-cycle income between workers with different innate abilities; and indirectly, to induce firms to absorb transient productivity shocks through efficiency wage contracts, leaving only the ability shock for the policy maker to handle. That is, the tax system induces the policy maker and firms to specialize in different shocks.

We begin with the case with only transient productivity shocks and take the labor income tax system as given. In this setting, we consider a long-term employment relationship between a firm-worker pair where the privateness of transient productivity shocks creates moral hazard. Importantly, we show that when labor income taxes are progressive, the firm incentivizes the worker with efficiency wage contracts that penalize terrible long-term performances rather than with short-term bonuses that vary with the concurrent output. Intuitively, the progressiveness of the tax system makes the after-tax consumption a concave function of the pre-tax income because as the latter increases, a larger portion of it is taxed away. Such concave relationship costs the firm a tax premium if she were to incentivize the worker with varying after-tax consumptions. In contrast, it increases the profitability of efficiency wage contracts which save the tax premium by paying the same after-tax consumption most of the time and provide near-optimal incentives by inflicting severe penalties for terrible long-term performances. In the limit case where transient productivity shocks occur infinitely frequently such that the long-term output becomes a perfect measure of the worker's long-term effort, incentive provision under the efficiency wage contract requires no consumption variation on the equilibrium path. Judging from the outcome, it is as if the firm absorbs all transient productivity shocks and avoids the tax premium all together.

We next build this insight into the original setting where the coexistence of persistent ability shocks and transient productivity shocks renders the policy complicated if the social planner were to implement the constrained social optimum by itself. Nevertheless, we manage to simplify the implementation strategy substantially by modeling explicitly the firm's interplay with the tax system. Specifically, we consider a game where the social planner proposes welfare policies before firms and workers form employment relationships through long-term contracts. In this setting, we propose an anonymous and time-invariant labor income tax sytsem that leaves firms zero profit if they offer efficiency wage contracts that specify the constrained optimal consumption-effort profiles. Under certain regularity conditions that guarantee the progressiveness of the tax system, such outcome is indeed an equilibrium of the market game. To see why, notice that the tax system is designed purposefully to make the planner's envelope condition and the firm's zero profit condition coincide. Consequently, from the viewpoint of a firm who wants to attract a particular ability type while facing the competition from the contracts offered to nearby types, the payoff that she needs to deliver — which is ultimately pinned down by the planner's envelope condition — always yields a non-positive profit. Since the efficiency wage contract yields zero profit, then it is necessarily optimal.

The rest of the paper proceeds as follows: Section 2 examines a setting with only transient productivity shocks and investigates firms' contractual responses to the labor income tax systsem; Section 3 studies an environment with both persistent and transient shocks and presents our core implementation strategy; Section 4 discusses the related literature; Section 5 concludes; Appendix A proposes an alternative formulation of the model in Section 2; Appendix B contains all the mathematical proofs.

2 Transient Productivity Shocks

Consider a long-term employment relationship between a firm-worker pair who interact infinitely frequently on a finite time interval $[0, 1]$ and take the labor income tax system $\{\tau(\cdot)\}$ as given. At $t = 0$, the firm proposes a contract which becomes binding if the worker prefers to participate than to consume the outside option that yields a flow utility $\underline{u}$. At each instant $t \in [0,1]$, the worker exerts a flow effort l_t at a cost $v(l_t)$ and produces a random output $\tilde{y}_t \in \{H, L\}$ with $\mathbb{P}(\tilde{y}_t = H|l_t) = p(l_t)$, where $v(\cdot)$ and $p(\cdot)$ satisfy $v(0) = 0, v'(\cdot) > 0, v''(\cdot) < 0$ and $p(\cdot) \in (0,1), p'(\cdot) > 0$ and $p''(\cdot) < 0$, respectively. The firm observes the output but not the effort. After y_t is realized, she pays a pre-tax wage w_t or effectively an after-tax consumption $c_t = w_t(1 - \tau(c_t))$ according to what the contract specifies.[1] Throughout this section we assume that $\tau(c)$ is strictly increasing in c such that as the pre-tax income increases, a larger fraction of it is taxed away and the corresponding after-tax consumption increases, too. The worker's flow payoff is $u(c_t) - v(l_t)$ where $u(\cdot) > 0$, $u''(\cdot) \leq 0$ and $\lim_{c \downarrow -\infty} u(c) = -\infty$.

In this setting, let us fix the effort l that the firm wants to elicit and compare two contracts: the *bonus contract* and the *efficiency wage contract*. To obtain the sharpest comparison, we assume that the worker is risk-neutral with $u(c) = c$ and leave it to the reader to verify that the result continues to hold when the worker is risk-averse.

Bonus Contract The bonus contract consists of a series of identical static contracts. At each instant t, it delivers an after-tax consumption $c(H)$ (or $c(L)$) if $y_t = H$ (or L). The contract is *incentive compatible* if it induces the worker to exert l at each instant, i.e.,

$$\mathbb{E}[c(\tilde{y})|l] - v(l) \geq \mathbb{E}[c(\tilde{y})|l'] - v(l') \tag{IC}$$

and is *individually rational* if makes the worker participate rather than consume the outside option, i.e.,

$$\mathbb{E}[c(\tilde{y})|l] - v(l) \geq \underline{u} \tag{IR}$$

[1]For notational convenience we write the tax rate as a function of the after-tax consumption rather than the pre-tax income.

The optimal bonus contract that elicit l solves

$$\max_{c(\cdot)} \mathbb{E}\left[\tilde{y} - \frac{c(\tilde{y})}{1-\tau(c(\tilde{y}))}\middle| l\right] \text{ s.t. (IC), (IR)}$$

We investigate the implication of labor income taxes in two steps. We first argue that to elicit a non-trivial effort, the optimal bonus contract should increase the after-tax consumption with the output to satisfy (IC) and set the average after-consumption to $v(l) + \underline{u}$ to make (IR) bind:

Lemma 1. *Take an arbitrary $l > 0$. The optimal bonus contract that elicits l satisfies $c(H) > c(L)$ and $\mathbb{E}[c(\tilde{y})|l] = v(l) + \underline{u}$.*

We next quantify the cost of providing incentives under progressive labor income tax systems. Specifically, let

$$\frac{v(l) + \underline{u}}{1 - \tau(v(l) + \underline{u})}$$

be the firm's pre-tax expenditure in the complete information benchmark with perfectly observable effort where it suffices to pay an after-tax consumption $v(l) + \underline{u}$ to make the worker participate. Define the *tax premium* as the difference between the pre-tax expenditure of the optimal bonus contract and its counterpart in the complete information benchmark, i.e.,

$$\text{Tax Premium} = \mathbb{E}\left[\frac{c(\tilde{y})}{1-\tau(c(\tilde{y}))}\middle| l\right] - \frac{v(l) + \underline{u}}{1 - \tau(v(l) + \underline{u})}$$

In the next proposition we show that the tax premium is always positive for any non-trivial target effort level. Intuitively, the progressiveness of the tax system means that as the pre-tax income increases, a larger fraction of it is taxed away, leaving the worker an after-tax consumption that is concave in the pre-tax income. This, together with the consumption variation that is necessary for incentive provision, makes the tax premium strictly positive—see Figure 2 for a graphical illustration. Formally,

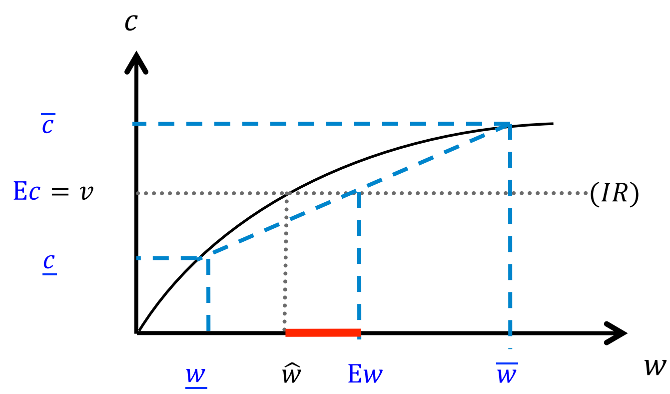

Figure 1: Tax Premium

Lemma 2. *Take an arbitrary $l > 0$. Under the optimal bonus contract that elicits l, the tax premium is strictly positive and the expected profit is strictly less than* $\mathbb{E}[\tilde{y}|l] - \frac{v(l)+\underline{u}}{1-\tau(v(l)+\underline{u})}$.

Efficiency Wage Contract The efficiency wage contract (c, μ, B) consists of a flow consumption c, a performance threshold μ and a penalty B. Specifically, it delivers a fixed after-tax consumption c at each $t \in [0, 1]$ and computes the sample average outputs $\int_0^1 y_t dt$ at $t = 1$. If the result exceeds the performance threshold μ, then the worker passes. Otherwise he fails and experiences a large utility loss B as a penalty.

Consider a special case where $c = v(l) + \underline{u}$, $\mu = \mathbb{E}[\tilde{y}|l]$ and $B > v(l)$. This case is interesting for two reasons. First, the contract induces an equilibrium where the worker exerts the target effort all the time. To see why, notice first that by using this strategy, the worker pass for sure according to the Law of the Large Numbers (see Sun (2006)). In fact, this is the cheapest way to pass because of the curvatures of

$p(\cdot)$ and $v(\cdot)$. Finally, observe that the worker prefers to pass than to fail, because in the case of failure, his total payoff is at most $-B$ where this upper bound is attained if he exerts no effort at all, and he is clearly better-off by exerting l all the time. Combindedly, these arguments establish that there exists a Bayesian Nash equilibrium of the efficiency wage contract where the worker always exerts l and passes for sure.

Second, it is noteworthy that in such equilibrium, the firm incurs the same pre-tax expenditure as in the complete information benchmark and pays no tax premium at all. This is because the efficiency wage contract pools information over time to obtain a very precise measure of the worker's long-term effort and relies on infrequent but severe penalties for incentive provision. In the continuous time limit, such mechanism is so powerful that incentive provision requires no consumption variation on the equilibrium path. This allows the firm to avoid the tax premium all together. Formally,

Lemma 3. *Take any arbitrary $l > 0$. Under the efficiency wage contract $(v(l) + \underline{u}, \mathbb{E}[\tilde{y}|l], B)$, there exists a Bayesian Nash equilibrium where the worker exerts l all the time and yields the same flow profit $\mathbb{E}[\tilde{y}|l] - \frac{v(l)+\underline{u}}{1-\tau(v(l)+\underline{u})}$ as in the complete information benchmark.*

Two remarks before we conclude this section. First, in the continuous-time limit, the firm absorbs all transient productivity shocks through efficiency wage contracts. Indeed, this is the key ingredient of our implementation strategy in the next section. Second, the results will remain qualitatively the same as long as we can form a precise estimate of the worker's performance within a given time span. In Appendix A.1 we formalizes this statement by analyzing a discrete-time finite-horizon model. There we show that while the effciency wage contract is no longer frictionless, it still provides near-optimal incentives at a negligible tax premium when the horizon is long.[2] For convenience we will use the continuous-time model as the building block of the subsequent analysis, but the reader should take the conclusions with a grain of salt.

3 Persistent and Transient Shocks

We now allow the worker to experience to both persistent ability shocks and transient productivity shocks. Specifically, suppse that at $t = 0$, the worker privately draws an ability type θ from a distribution $F(\cdot)$ with support $\Theta = [\underline{\theta}, \infty)$. At each $t \in [0, 1]$, he exerts a flow effort l_t and yields the same output distribution as before, but his flow payoff is now $u(c_t) - v(l_t, \theta)$ where $v_l > 0, v_{ll} > 0, v_\theta < 0$ and $v_{l\theta} < 0$.

We proceed in two steps. We first derive the constrained optimal allocation from a relax program of the social planner and argue that it requires a personalized and history-dependent labor income tax system for the planner to implement this allocation by itself. We next argue that such allocation is indeed implementable by a simple progressive labor income tax system when we introduce firms to the framework and model explictly their contractual responses the labor income tax sytem.

3.1 The Planner's Problem

For the time being, ignore the firm and consider the problem of a social planner who wants to maximize the weighted life cycle utilities of all types of workers (use $\lambda(\theta)$ to denote the Pareto weight of type θ worker), subject to the information constraint that

[2]We also establish that the performance of the efficiency wage contract is robust to contracutal frictions such as private saving, discounting and limited liability in a certain range of parameters.

it observes effort, the output but not the ability type,[3] and the resource constraint that it must collect a revenue R over $[0,1]$. Given the stationarity of the production technology and the curvature of worker's utility functions, the problem boils down to solving the menu of consumption-effort profiles $\{c(\theta), l(\theta)\}$ that maximizes the weighted flow utility of all types of workers, subject to the truth-telling constraint that each type of worker selects the profile that matches his true ability type and the resource constraint that the flow revenue is at least R, i.e.,

$$\begin{aligned}
\max_{\{c(\theta),l(\theta)\}} \int_{\underline{\theta}}^{\infty} & \lambda(\theta)(c(\theta) - v(l(\theta),\theta))dF(\theta) \\
\text{s.t. } & u(c(\theta)) - v(l(\theta),\theta) \geq u(c(\theta')) - v(l(\theta'),\theta), \ \forall \theta, \theta' \\
\int_{\underline{\theta}}^{\infty} & \mathbb{E}[\tilde{y}|l(\theta)] - c(\theta)dF(\theta) \geq R
\end{aligned}$$

This is a standard screening problem and can be solved by applying the Envelope Theorem of Milgrom and Segal (2002). We characterize the solution $\{c^*(\theta), l^*(\theta)\}$ in the next lemma—readers who are interested in the details should consult Appendix B for the proof:

Lemma 4. $\{c^*(\theta), l^*(\theta)\}$ *satisfies*

$$u(c^*(\theta)) = v(l^*(\theta),\theta) - \int_{\theta}^{\infty} v_\theta(l^*(s),s)ds + U(\underline{\theta}) \tag{ICFOC}$$

where $U(\underline{\theta})$ *is the utility of type* $\underline{\theta}$ *worker, and*

$$\frac{\int_{\theta}^{\infty}(\lambda(s)-\mu)dF(s)}{f(\theta)} v_{l\theta}(l^*(\theta),\theta) = \mu\left[p'\left(l^*(\theta)\right)(H-L) - v_l\left(l^*(\theta),\theta\right)\right]$$

[3] Again, the result of this relaxed program is an upper bound for what could be implemented. As we will see soon, such result is indeed implementable.

for some $\mu > 0$.

The central research question of this paper concerns the implementation of $\{c^*(\theta), l^*(\theta)\}$. Before we present our core implementation strategy, it is useful to observe that if the social planner were to implement this allocation by itself through labor income taxes, then it should adopt a personalized tax system that varies the tax rate from person to person and shock to shock such that two workers with different ability types θ, θ' are subject to different tax rates $\frac{y-c^*(\theta)}{y}, \frac{y-c^*(\theta')}{y}$ even when they produce the same output y. Such solution is rather complicated and will face a lot of challenges when it comes to implementing it in reality.

3.2 Implementation with Firms

We now present our core implementation strategy. Consider a game involving a social planner, a continuum of firms indexed by $i \in [0, 1]$ who observe only the output but not the ability type or the effort, and a continuum of workers whose ability types are drawn i.i.d. from $F(\cdot)$. Time evolves as follows. At $t = 0$, the social planner specifies an anonymous and time-invariant labor income tax system $\{\tau(\cdot)\}$. Then each firm i proposes a menu of contracts $\{\{c_{i,t}(\cdot;\theta), l_{i,t}(\cdot;\theta)\}_t\}_\theta$ that consists of consumption-effort plans[4] for all types of workers, and each worker opts into the contract that maximizes his expected life cycle utility. Then the contract becomes binding[5] and production takes place on $[0, 1]$.

We say that $\{\tau(\cdot)\}$ *implements* $\{\{\{c_{i,t}(\cdot;\theta), l_{i,t}(\cdot;\theta)\}_t\}_\theta\}_i$ *if there exists a Bayesian Nash equilibrium where*

- Each firm i finds $\{\{c_{i,t}(\cdot;\theta), l_{i,t}(\cdot;\theta)\}_t\}_\theta$ optimal given other firms' contracts

[4] These plans depend on the output history which is written as "$\cdot$" for notational convenience.

[5] This is not a restrictive assumption. Indeed, the result continues to hold even if we allow for recontracting after the worker's type and output history is publicly revealed.

$\{\{\{c_{j,t}(\cdot;\theta), l_{j,t}(\cdot,\theta)\}_t\}_\theta\}_{j\neq i}$ and workers' strategies;

- Each type θ worker opts into $\{c_{i,t}(\cdot;\theta), l_{i,t}(\cdot;\theta)\}$ for some i and exerts the recommended effort $\{l_{i,t}(\cdot;\theta)\}$ on $[0,1]$.

In words, this means that the tax system induces an equilibrium where each firm proposes the optimal contract menu given other firms' contracts and workers' strategies, whereas each worker opts into the optimal contract that matches his ability type and exerts the recommended effort throughout his life cycle.

Consider a tax system $\{\tau^*(\cdot)\}$ defined as follows:

$$\tau^*(c) = 1 - \frac{c^*(\theta)}{\mathbb{E}[\tilde{y}|l^*(\theta)]} \text{ if } c = c^*(\theta) \tag{T}$$

Intuitively, such system yields zero profit to all firms if it implements an outcome where the contract attracting each type θ worker is an efficiency wage contract that specifies the constrained optimal consumption-effort profile $(c^*(\theta), l^*(\theta))$. The main result of this paper asserts that such outcome is indeed an equilibrium provided that certain regularity conditions that guarantee the progressiveness of $\{\tau^*(\cdot)\}$ are satisfied. Formally,

Theorem 1. *Suppose that* $\frac{\mathbb{E}[\tilde{y}|l^*(\theta)]}{c^*(\theta)} < \frac{p'(l^*(\theta))u'(c^*(\theta))(H-L)}{v_l(l^*(\theta),\theta)}$. *Then* $\{\tau^*(\cdot)\}$ *is progressive and implements the following outcome:*

(i) Each firm i *offers an efficiency wage contract* $(c^*(\theta), \mathbb{E}[\tilde{y}|l^*(\theta)], B(\theta))$ *to a certain ability type* θ *and a contract that yields a lower life-cycle utility than* $(c^*(\theta'), \mathbb{E}[\tilde{y}|l^*(\theta')], B(\theta'))$ *to any other ability type* θ'*;*

(ii) Type θ *worker self-selects into* $(c^*(\theta), \mathbb{E}[\tilde{y}|l^*(\theta)], B(\theta))$ *and exerts* $l^*(\theta)$ *everywhere on* $[0,1]$ *for all* θ*;*

(iii) All firms earn zero flow profit everywhere on $[0,1]$.

Figure 3.2 illustrates the key idea behind Theorem 1, where the worker's indifference curve is colored in blue whereas the envelope condition (i.e., (ICFOC)) of the planner's problem and the firm's zero profit curve are colored in red. Crucially, the tax system is designed purposefully to make the last two curves coincide. Albeit simple, this observation has important implications for the firm's equilibrium behavior. To see why, let us consider the problem of firm i who wants to attract type θ worker. Clearly, she can do so by offering $(c^*(\theta), \mathbb{E}[\tilde{y}|l^*(\theta)], B(\theta))$ and earn zero profit, but can she do better? The answer is negative. To see why, notice that to attract θ, the payoff that i needs to deliver is pinned down by the equilibrium payoffs of nearby types which, by construction, are determined by the envelope condition of the planner's problem. Then it follows the tangency between θ's indifference curve and the envelope curve (and thus the zero profit curve) that any attempt to extract a positive profit from θ will cause him to opt into the contracts offered to nearby types.

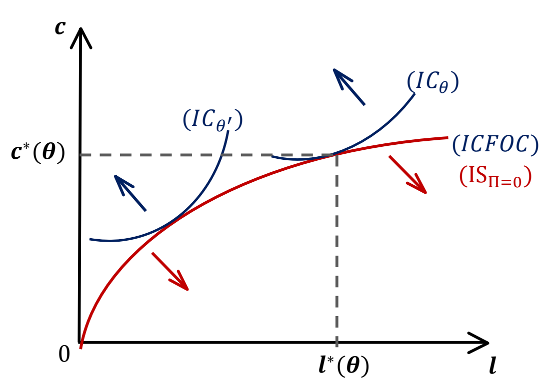

Figure 2: Implementation with Firms

An immediate corollary of the above argument is that $\{\tau^*(\cdot)\}$ continues to implement the outcome stated in Theorem 1 even if after the initial contracting stage, workers' ability types become public information parties are allowed to recontract at any time. Again, this is because the tax system and the interaction between firms at the initial contracting stage have already driven the flow profit from each ability types to zero.

$\{\tau^*(\cdot)\}$ is progressive if and only if $\frac{\mathbb{E}[\tilde{y}|l^*(\theta)]}{c^*(\theta)} < \frac{p'(l^*(\theta))u'(c^*(\theta))(H-L)}{v_l(l^*(\theta),\theta)}$. To obtain a heuristic interpretation of this condition, let us restrict attention to the case of risk-neutral workers such that $u'(c) = 1$, where it follows by Lemma 4 that the previous condition is equivalent to

$$\frac{\mathbb{E}[\tilde{y}|l(\theta)]}{v(l(\theta),\theta) - \int_{\underline{\theta}}^{\theta} v_\theta(l(s),s)ds} < 1 + \frac{\int_\theta^\infty (\mu - \lambda(s))dF(s)}{\mu f(\theta)} \frac{|v_{l\theta}(l(\theta),\theta)|}{v_l(l(\theta),\theta)}$$

For this condition to hold, we need the right hand side to be large, which happens if $\lambda(\theta)$ is small for large θ and (or) if $|v_{l\theta}|$ is large. Intuitively, the first situation happens if the social planner does not care too much about high ability workers and therefore is willing to tax them heavily, whereas the second situation occurs if the complementarity between ability and effort is strong enough such that high ability workers still prefer high-effort-high-pay jobs despite that they are being taxed heavily.

To summarize, the progressive labor income tax system in this section plays two roles: to redistribte the life-cycle income between workers of different ability types, and to induce firms to absorb transient productivity shocks through efficiency wage contracts. It is carefully designed such that in the resulting market equilibrium, each firms optimally hires a certain ability type and all firms earn zero expected flow profit. By exploiting the interaction between the tax system, the firm's contractual choice and the interaction between firms, we manage to implement the constrained social optimum by a simple welfare policy that is commonly observed in practice.

4 Related Literature

5 Conclusion

We use an anonymous and time-invariant progressive labor income tax system to implement the constrained social optimum in a setting where workers privately experience a persistent ability shock and many transient productivity shocks throughout their life cycles. Our implementation strategy builds upon the interplay between the tax system and firm's contracutal choices. In particular, we show that when labor income taxes are progressive, firms overcome the moral hazard problem caused by transient productivity shocks with efficiency wage contracts which — in the limit case where shocks are infinitely frequent — completely absorb these shocks on the equilibrium path, leaving only the persistent ability shock for the policy maker to handle. Based on this insight, we design an anonymous and time-invariant progressive income tax system to attain the socially optimal redistribution.

A Appendix: Alternative Formulation

A.1 Long-Term Contracting with Moral Hazard in a Discrete Time-Finite Horizon Model

In this section we establish the counterpart of Lemma 3 in a discrete-time finite-horizon model with moral hazard.[6] The employment relationships lasts for T periods, $t = 1, 2, ..., T$, and there is no discounting. The worker has the same utility function and has access to the same production technology as before, where $p(\cdot), u(\cdot), v(\cdot)$ are

[6]The argument in this section continues to hold when the worker privately observes the productivity shock before choosing the effort.

assumed to be smooth in their arguments. The firm observes the output but not the effort. At the outset, she takes the progressive labor income tax system $\{\tau(\cdot)\}$ as given and proposes an efficiency wage contract $(c_T, \mu(l) - b_T, B_T)$ which becomes binding if the worker prefers to participate than to consume his outside option that yields a per-period utility $\underline{u}$.

The efficiency wage contract delivers a fixed after-tax consumption c_T in period $t = 1, 2, \cdots, T$ and computes the sample mean of outputs mean of outputs $\mu_T = \frac{1}{T}\sum_{t=1}^{T} y_t$ at the end of period T. If the result exceeds the threshold $\mu(l) - b_T$ where $\mu(l) \doteq \mathbb{E}[\tilde{y}|l]$ and $b_T \sim \mathcal{O}(T^{-\frac{1}{2}+\varepsilon})$ for some arbitrary $\varepsilon \in \left(0, \frac{1}{2}\right)$, then the worker passes and nothing happens. Otherwise he fails and experiences a utility loss $B_T = \alpha v(l) T$ for some arbitrary $\alpha > 1$.

The contract induces a dynamic game, where $y^t = (y_1, \cdots, y_t)$ and $l^t = (l_1, \cdots, l_t)$ denote a t-period output history and a t-period effort profile, respectively. Without loss of generality, let us restrict attention to public strategies $\sigma = \{\sigma_t : \{H, L\}^{t-1} \rightarrow \mathbb{R}_+\}_{t=1}^{T}$ where $\sigma_t(\cdot)$ maps the $(t-1)$-period output history to the period-t effort choice. The solution concept is Bayesian Nash equilibrium.

The main result of this section states that when the horizon is long, the efficiency wage contract yields a per-period expected profit that is arbitrarily close to its counterpart in the complete information benchmark. Formally,

Theorem 2. *Take an arbitrary l. When T is large, there exists a Bayesian Nash equilibrium σ_T^* of the efficiency wage contract $(c_T, \mu(l) - b_T, B_T)$ where*

(i) The worker rarely fails: $\mathbb{P}^{\sigma_T^*}\left(\mu_T < \mu(l) - b_T\right) \sim \mathcal{O}(b_T)$*;*

(ii) The per-period expected effort cost is close to $v(l)$: $|\mathbb{E}^{\sigma_T^*}[\frac{1}{T}\sum_{t=1}^{T} v(l_t)] - v(l)| \sim \mathcal{O}(b_T)$*;*

(iii) The per-period expected profit is close to its counterpart in the complete information benchmark:

$$\left|\frac{1}{T}\mathbb{E}^{\sigma_T^*}\left[\sum_{t=1}^{T}\tilde{y}_t-\frac{c_t}{1-\tau(c_t)}\right]-\mathbb{E}\left[\tilde{y}-\frac{u^{-1}(v(l)+\underline{u})}{1-\tau(u^{-1}(v(l)+\underline{u}))}\middle|l\right]\right|\sim\mathcal{O}(b_T)$$

The key to the proof of Theorem 2 is to define a critical event and to characterize its necessary properties. Specifically, for any T-period effort profile l^T, define $\tilde{\xi}_t(l_t)=\tilde{y}_t-\mu(l_t)$ as the period-t productivity shock and $\mathcal{E}(l^T)$ as the event where the sample mean of productivity shocks are bounded around its theoretical mean 0 by b_T, i.e.,

$$\mathcal{E}(l^T)=\left\{(\xi_1(l_1),\cdots,\xi_t(l_t)):\left|\frac{1}{T}\sum_{t=1}^{T}\xi_t(l_t)\right|\leq b_T\right\}$$

We make two observations about $\mathcal{E}(l^T)$. First, since $\{\tilde{\xi}_t(l_t)\}$ are bounded independent random variables, it follows by McDiarmid (1989)'s concentration inequality that there is a uniform upper bound for the probability of $\mathcal{E}^c(l^T)$ that shrinks exponentially fast as T increases:

Lemma 5. *Take an arbitrary l^T. Then*

(i) $\mathbb{P}\left(\mathcal{E}(l^T)\right)\geq 1-2\exp\left(-\frac{2Tb_T^2}{(H-L)^2}\right)$ *for every T;*

(ii) $\mathbb{P}(\mathcal{E}^c(l^T))\sim o(b_T)$ *as* $T\to\infty$.

Second, if the worker passes the test at $\mathcal{E}(l^T)$, then the total effort cost must be sufficiently close to $v(l)T$ because otherwise there exists a mixed effort profile that yields more or less the same revenue as the target effort l and yet is significantly cheaper, contradicting the smoothness assumption on the production technology. Formally,

Lemma 6. *At $\mathcal{E}(l^T)$, if $\mu_T\geq\mu(l)-b_T$, then $\frac{1}{T}\sum_{t=1}^{T}v(l_t)\sim v(l)(1-\mathcal{O}(b_T))$ when T is sufficiently large.*

Proof. Proof by contradiction. Define $\hat{l} = \frac{1}{T}\sum_{t=1}^{T} l_t$. Then it follows by the curvature of $v(\cdot)$ and $p(\cdot)$ that

$$\mu\left(\hat{l}\right) \geq \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} \tilde{y}_t | l^T\right] \text{ and } v\left(\hat{l}\right) \leq \frac{1}{T}\sum_{t=1}^{T} v(l_t)$$

Now suppose to the contrary that $\frac{1}{T}\sum_{t=1}^{T} v(l_t) \ll v(l) - \mathcal{O}(b_T)$ when T is large. Then it follows by the assumption $\mu_T \geq \mu(l) - b_T$ that

$$\frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T} \tilde{y}_t | l^T\right] \geq \mathbb{P}\left(\mathcal{E}\left(l^T\right)\right)\left(\mu(l) - b_T\right) + \mathbb{P}\left(\mathcal{E}^c\left(l^T\right)\right) L = \mu(l) - \mathcal{O}(b_T)$$

and hence that $\hat{l}$ satisfies

$$\mu\left(\hat{l}\right) = \mu(l) - \mathcal{O}(b_T) \text{ and } v\left(\hat{l}\right) \ll v(l) - \mathcal{O}(b_T)$$

Now since both $\mu(\cdot)$ and $v(\cdot)$ are smooth, then the fact that

$$l - \hat{l} \approx \frac{\mu(l) - \mu\left(\hat{l}\right)}{\mu'(l)} \sim \mathcal{O}(b_T)$$

implies that

$$v(l) - v\left(\hat{l}\right) \approx v'(l)\left(l - \hat{l}\right) \sim \mathcal{O}(b_T)$$

a contradiction. □

Given Lemma 5 and 6, we now state the proof of Theorem 2:

Proof. Part (i): use $\mathcal{F}$ to denote the event where the worker fails the test. Define

$$\pi_T(\mathcal{F}) = \sum_{l^T} \mathbb{P}^{\sigma_T^*}(l^T)\mathbb{P}\left(\mathcal{E}\left(l^T\right)\right)\mathbb{P}\left(\mathcal{F}|l^T, \mathcal{E}\left(l^T\right)\right)$$
$$\pi_T(\mathcal{F}^c) = \sum_{l^T} \mathbb{P}^{\sigma_T^*}(l^T)\mathbb{P}\left(\mathcal{E}\left(l^T\right)\right)\mathbb{P}\left(\mathcal{F}^c|l^T, \mathcal{E}\left(l^T\right)\right)$$

and notice that $1 - \pi_T(\mathcal{F}) - \pi_T(\mathcal{F}^c) \leq 2\exp\left(-\frac{2Tb_T^2}{(H-L)^2}\right) \sim o(b_T)$.

In equilibrium, the worker's expected payoff is bounded from above by

$$Tu(c_T) - \left[\pi_T(\mathcal{F})B_T + \pi_T(\mathcal{F}^c)(v(l)T - \mathcal{O}(Tb_T))\right]$$

Now suppose that he exerts the target effort l all the time instead, then his expected payoff is bounded from below by

$$Tu(c_T) - Tv(l) - 2\exp\left(-\frac{2Tb_T^2}{(H-L)^2}\right)B_T$$

Subtracting the second expression from the first one yields an upper bound for the net benefit of the equilibrium strategy:

$$Tv(l)\left[-(\alpha - 1)\pi_T(\mathcal{F}) + o(b_T)\right] + \mathcal{O}(Tb_T)$$

In particular, if $\pi_T(\mathcal{F}) \gg \mathcal{O}(b_T)$, then the above expression is strictly negative when T is large, implying that the worker is strictly better-off by exerting l all the time, a contradiction. Now that $\pi_T(\mathcal{F}) \sim \mathcal{O}(b_T)$, we can bound the equilibrium probability of failure from above by

$$\mathbb{P}^{\sigma_T^*}(\mathcal{F}) \leq \pi_T(\mathcal{F}) + \sum_{l^T} \mathbb{P}^{\sigma_T^*}(l^T)\mathbb{P}\left(\mathcal{E}^c\left(l^T\right)\right) \sim \mathcal{O}(b_T) + o(b_T) \sim \mathcal{O}(b_T)$$

Part (ii): first, bound the equilibrium per-period effort cost $\frac{1}{T}\mathbb{E}^{\sigma_T^*}\left[\sum_{t=1}^{T} v(l_T)\right]$ from below by

$$\pi_T(\mathcal{F}^c)(v(l) - \mathcal{O}(b_T)) \sim v(l) - \mathcal{O}(b_T)$$

Second, notice that if this cost exceeds $v(l) + \mathcal{O}(b_T)$, then the worker's equilibrium per-period payoff is less than $u(c_T) - v(l) - \mathcal{O}(b_T)$ and he is better-off by exerting l all the time, a contradiction. Therefore, we have

$$\left|\frac{1}{T}\mathbb{E}^{\sigma_T^*}\left[\sum_{t=1}^{T} v(l_T)\right] - v(l)\right| \sim \mathcal{O}(b_T)$$

Part (iii): since the worker's equilibrium per-period payoff is bounded from below by $u(c_T) - v(l) - \mathcal{O}(b_T)$, it suffices to pay $c_T \sim u^{-1}(v(l) + \underline{u} + \mathcal{O}(b_T))$ to satisfy his ex-ante participation constraint. Therefore the firm's equilibrium per-period profit is bounded from below by

$$\pi_T(\mathcal{F}^c)(\mu(l) - b_T) + (1 - \pi_T(\mathcal{F}^c))L - \frac{c_T}{1 - \tau(c_T)} \sim \mu(l) - \frac{u^{-1}(v(l) + \underline{u})}{1 - \tau\left(u^{-1}(v(l) + \underline{u})\right)} - \mathcal{O}(b_T)$$

$\square$

A.2 Private Saving

We show that when the worker has access to private savings, the efficiency wage contract continues to provide near-optimal incentives and to absorb most transient productivity shocks provided that the worker's desire for consumption smoothing is moderate. The results suggests that private market may still attain near-efficient outcomes even in the presence of unobservable trades.

Formally, let us modify the setting in the previous section by allowing the worker to store the after-tax consumption at zero interest rate and consume it at any time in the future. Then the efficiency wage contract induces a dynamic game where a typical t-period history of efforts, outputs and consumptions are denoted by l^t, y^t, ψ^t, respectively. The worker's strategy is now $\sigma = \{l_t(\cdot), \psi_t(\cdot)\}$, where $l_t : \mathbb{R}_+^{t-1} \times \{H, L\}^{t-1} \times \mathbb{R}^{t-1} \to \mathbb{R}_+$ maps the $(t-1)$-period histories of efforts, outputs and consumptions to the period-t effort, whereas $\psi_t : \mathbb{R}_+^t \times \{H, L\}^t \times \mathbb{R}^{t-1} \to \mathbb{R}$ maps the t-period histories of efforts and outputs and the $(t-1)$-period history of consumptions into the the period-t consumption, subject to the budget constraint that $\sum_{s=1}^t \psi_t \leq \sum_{s=1}^t c_s$ for all t and (l^t, y^t, ψ^t).

In this setting, a slight modification of Theorem 2 yields the following performance bound for the efficiency wage contract:

Corollary 1. *Consider an efficienc wage contracts $(c_T, \mu(l) - b_T, B_T)$ where $b_T \sim \mathcal{O}(T^{-\frac{1}{2}+\varepsilon})$ for some arbitrary $\varepsilon \in \left(0, \frac{1}{2}\right)$ and $B_T > v(l)T$. When T is large, if (c_T, B_T) satisfy*

(a) $u(c_T) \geq v(l) + \underline{u} + \mathcal{O}(b_T)\left(1 + \frac{B_T}{T}\right)$;

(b) $u(c_T) - u\left(\frac{(T-1)c_T + u^{-1}(u(c_T) - B_T)}{T}\right) = \alpha v(l)$ *for some arbitrary* $\alpha > 1$;

(c) $\exp\left(-\frac{2Tb_T^2}{(H-L)^2}\right)\frac{B_T}{T} \sim \mathcal{O}(b_T)$

then there exists a Bayesian Nash equilibrium σ_T^* where

(i) The worker rarely fails: $\mathbb{P}^{\sigma_T^*}(\mathcal{F}) \sim \mathcal{O}(b_T)$;

(ii) The per-period effort cost is close to $v(l)$: $\left|\frac{1}{T}\mathbb{E}^{\sigma_T^*}\left[\sum_{t=1}^T v(l_T)\right] - v(l)\right| \sim \mathcal{O}(b_T)$;

(iii) The worker is willing to participate: $\frac{1}{T}\mathbb{E}^{\sigma_T^*}[u(c_t) - v(l_t)] \geq \underline{u}$;

(iv) The per-period profit is at least $\mu(l) - \frac{c_T}{1-\tau(c_T)} - \mathcal{O}(b_T)$.

In particular, if $\lim_{T\to\infty} c_T = v(l) + \underline{u}$, *then the contract yields a near-optimal profit* $\lim_{T\to\infty} \frac{1}{T}\mathbb{E}^{\sigma_T^*}\left[\tilde{y}_t - \frac{c_t}{1-\tau(c_t)}\right] = \mathbb{E}[\tilde{y}|l] - \frac{u^{-1}(v(l)+\underline{u})}{1-\tau(u^{-1}(v(l)+\underline{u})}$

Proof. Part (i): the worker's equilibrium per-period payoff is now bounded from above by

$$(1-\pi_T(\mathcal{F}))u(c_T) - \pi_T(\mathcal{F}^c)[v(l) - \mathcal{O}(b_T)] + \pi_T(\mathcal{F})u\left(\frac{(T-1)c_T + u^{-1}(u(c_T) - B_T)}{T}\right)$$

whereas his per-period payoff from exerting l all the time and saving nothing is bounded from below by

$$u(c_T) - v(l) - 2\exp\left(-\frac{2Tb_T^2}{(H-L)^2}\right)\frac{B_T}{T}$$

By assumption, the net benefit of the equilibrium strategy is bounded from above by

$$\pi_T(\mathcal{F})\left[u\left(\frac{(T-1)c_T + u^{-1}(u(c_T) - B_T)}{T}\right) - u(c_T) + v(l)\right] + 2\exp\left(-\frac{2Tb_T^2}{(H-L)^2}\right)\frac{B_T}{T} + \mathcal{O}(b_T)$$
$$\sim -\pi_T(\mathcal{F})(\alpha-1)v(l) + \mathcal{O}(b_T)$$

Thus, if $\pi_T \gg \mathcal{O}(b_T)$, then the above expression is negative and the worker strictly prefers the deviation strategy to the equilibrium strategy, a contradiction. Therefore, we have $\mathbb{P}^{\sigma_T^*}(\mathcal{F}) = \pi_T(\mathcal{F}) + o(b_T) \sim \mathcal{O}(b_T)$.

Part (ii): follows by Theorem 2 (ii).

Part (iii): the worker's per-period equilibrium payoff is bounded from below by

$$u(c_T) - (v(l) + \mathcal{O}(b_T)) - \mathcal{O}(b_T)\frac{B_T}{T} = u(c_T) - v(l) + \mathcal{O}(b_T)\left(1 + \frac{B_T}{T}\right)$$

By assumption, the RHS is greater than $\underline{u}$ and thus the worker's ex-ante participation constraint is satisfied.

Part (iv): obvious. $\square$

Condition (b) illustrates how private saving increases the cost of incentive provision. To see why, notice that the LHS is bounded from below by

$$\begin{aligned}
& u(c_T) - u\left(\frac{(T-1)c_T + u^{-1}(u(c_T) - B_T)}{T}\right) \\
&= u(c_T) - u\left(c_T - \frac{u^{-1}(u(c_T)) - u^{-1}(u(c_T) - B_T)}{T}\right) \\
&\geq \underbrace{\frac{u'(c_T)}{u'(u^{-1}(u(c_T) - B_T))}}_{<1, \uparrow \text{ in } c_T} \frac{B_T}{T}
\end{aligned}$$

where the coefficient in front of $\frac{B_T}{T}$ is strictly less than one and is increasing in c_T. This implies that Condition (b) is (1) less likely to hold than its counterpart when there is no saving whereby the per-period utility of consumption between success and failure differ simply by $\frac{B_T}{T}$, and (2) is more likely to hold when c_T is large such that the difference in the marginal utility of consumption between success and failure is small. The intuition behind this result is straightforward: saving destroys the incentive to supply effort because it makes failure less painful; to discourage saving at the lowest cost, the firm increases the baseline consumption in the efficiency wage contract until the deterrent power of the penalty is restored.

A.3 Other Contractual Frictions

Certainly, the performance of the efficiency wage contract is undermined by contractual frictions such as discounting and limited liability. If workers care little about the future because of job mobility or macroeconomic uncertainty, or if they cannot be severely penalized because of the bankruptcy law or a malfunctioning court system, then the firm's capacity of penalizing terrible long-term performances is severely limited. A partial remedy to this problem is to divide the employment relationship into relatively short blocks and to apply the efficiency wage contract in each block. Nevertheless, such remedy comes at a cost because the precision of the statistical test decreases as the observations per block shrinks, resulting in an increase in type I error whose welfare implication is beyond the scope of the current analysis. The relevance of contractual frictions depends the nature of the job, too. For example, if the job generates many data points on employee performance within a short time span, then the aforementioned constraints are not as binding as they seem.

B Appendix: Omitted Proofs

Proof of Lemma 1

Proof. (IC) implies that $(p(l) - p(l')(c(H) - c(L))) \geq v(l) - v(l')$ for all l, l'. Since $p(\cdot), v(\cdot)$ are strictly increasing, then we must have $c(H) > c(L)$.

Meanwhile, (IR) must bind under the optimal contract, i.e., $\mathbb{E}[c(\tilde{y})|l] = v(l) + \underline{u}$, because otherwise there exists $\varepsilon > 0$ such that $\{c(H) - \varepsilon, c(L) - \varepsilon\}$ yields a strictly higher profit without violating (IC) and (IR). $\square$

Proof of Lemma 2

Proof. By assumption, $\frac{c}{1-\tau(c)}$ is strictly convex in c. Then it follows by Lemma 1 that

$$\mathbb{E}\left[\frac{c(\tilde{y})}{1-\tau(c(\tilde{y}))}\middle| l\right] > \frac{\mathbb{E}[c(\tilde{y})|l]}{1-\tau(\mathbb{E}[c(\tilde{y})|l])} = \frac{v(l)+\underline{u}}{1-\tau(v(l)+\underline{u})}$$

$\square$

Proof of Lemma 3

Proof. See the proof sketch in the main body of the paper. $\square$

Proof of Lemma 4

Proof. It follows by Milgrom and Segal (2002) that $\{c^*(\theta), l^*(\theta)\}$ satisfies the truth-telling constraint of ability if and only if

$$u(c^*(\theta)) = v(l^*(\theta), \theta) - \int_{\theta}^{\infty} v_\theta(l^*(s), s)ds + U(\underline{\theta}) \tag{ICFOC}$$

where $U(\underline{\theta})$ denotes the reservation utility of the lowest type worker. Thus, rewrite the planner's problem as

$$\max_{\{c(\theta), l(\theta)\}} \int_{\underline{\theta}}^{\infty} \lambda(\theta)(c(\theta) - v(l(\theta), \theta))dF(\theta)$$
$$\text{s.t. (ICFOC) and } \int_{\underline{\theta}}^{\infty} \mathbb{E}[\tilde{y}|l(\theta)] - c(\theta)dF(\theta) \geq R$$

Let μ be the Lagrangian multiplier of the revenue constraint and transform this problem into the following unconstrained optimization problem:

$$\max_{\{l(\theta)\}} - \int_{\underline{\theta}}^{\infty} \lambda(\theta) \int_{\underline{\theta}}^{\theta} v_\theta(l(s), s)dsdF(\theta) + \mu \int_{\underline{\theta}}^{\infty} \mathbb{E}[\tilde{y}|l(\theta)] - v(l(\theta), \theta) + \int_{\underline{\theta}}^{\theta} v_\theta(l(s), s)dsdF(\theta)$$

Apply Fubini's Theorem and further simplify the objective function to

$$\max_{\{l(\theta)\}_{\theta=\underline{\theta}}^{\theta=\infty}} \int_{\underline{\theta}}^{\infty} -\Lambda(\theta)v_\theta(l(\theta), \theta) + \mu \left(\mathbb{E}[\tilde{y}|l(\theta)] - v(l(\theta), \theta) + \frac{1 - F(\theta)}{f(\theta)} v_\theta(l(\theta), \theta) \right) dF(\theta)$$

where $\Lambda(\theta) = \frac{\int_{\theta}^{\infty} \lambda(s)dF(s)}{f(\theta)}$. Finally, pointwise optimization with respect to $l(\theta)$ yields the result. $\square$

Proof of Theorem 1

Proof. First, we assume that $\{\tau^*(\cdot)\}$ is progressive such that firms will optimally offer efficiency wage contracts. In this setting, we verify that the outcome stated in Theorem 1 is indeed an equilibrium. The proof has three steps:

- First, it follows by (ICFOC) and (T) that type θ worker will indeed opt into $(c^*(\theta), \mathbb{E}[\tilde{y}|l^*(\theta)], B(\theta))$ and exert $l^*(\theta)$ at all $t \in [0,1]$.

- Second, given the equilibrium contracts offered to $\theta' \neq \theta$, the equilibrium payoff of type θ worker is at least

$$
\begin{aligned}
&\lim_{\theta' \to \theta} u\left(c^*(\theta')\right) - v\left(l^*(\theta'), \theta\right) \\
=&\lim_{\theta' \to \theta} u\left(c^*(\theta')\right) - v\left(l^*(\theta'), \theta'\right) + \lim_{\theta' \to \theta} \int_{\theta}^{\theta'} v_\theta\left(l^*(\theta'), s\right) ds \\
=&\lim_{\theta' \to \theta} \left[U(\underline{\theta}) - \int_{\underline{\theta}}^{\theta'} v_\theta(l^*(s), s) ds \right] \\
=&u(c^*(\theta)) - v(l^*(\theta), \theta)
\end{aligned}
$$

  Based on this observation, we argue that the equilibrium efficiency wage contract that attracts θ is $(c^*(\theta), \mathbb{E}[\tilde{y}|l^*(\theta)], B(\theta))$. To see why, notice that first that such contract does attract θ and yield zero profit. To establish its optimality, compare it with a different efficiency wage contract $(c, \mathbb{E}[\tilde{y}|l], B)$ that yields the same payoff to θ such that

$$
u(c) - v(l, \theta) = u(c^*(\theta)) - v(l^*(\theta), \theta) = -\int_{\underline{\theta}}^{\theta} v_\theta(l^*(s), s) ds
$$

  Define $\theta(l)$ by $l^*(\theta(l)) = l$ and let $c(l) = c^*(\theta(l))$. Then we have

$$
u(c(l)) - v(l, \theta(l)) = -\int_{\underline{\theta}}^{\theta(l)} v_\theta(l^*(s), s) ds
$$

Subtracting the second equation from the first one yields

$$\begin{aligned}u(c)-u(c(l)) =& v(l,\theta)-v(l,\theta(l))+\int_{\theta}^{\theta(l)} v_\theta(l^*(s),s)ds \\ =& -\int_{\theta}^{\theta(l)} v_\theta(l(s),s)ds+\int_{\theta}^{\theta(l)} v_\theta(l^*(s),s)ds \\ =& \int_{\theta}^{\theta(l)} v_\theta(l^*(s),s)-v_\theta(l,s)ds\end{aligned}$$

where the assumption $v_{\theta l} < 0$ (which implies that $l^*(\cdot)$ is strictly increasing) implies that the last line is strictly negative as long as $l \neq l^*(\theta)$. Consequently,

$$\mathbb{E}[\tilde{y}|l]-\frac{c}{1-\tau(c)} < \mathbb{E}[\tilde{y}|l]-\frac{c(l)}{1-\tau(c(l))}$$

That is, $(c, \mathbb{E}[\tilde{y}|l], B)$ yields a strictly lower profit than $(c^*(\theta), \mathbb{E}[\tilde{y}|l^*(\theta)], B(\theta))$.

- Third, observe that firms are indifferent between all ability types because they all yield zero flow profit.

Second, we provide sufficient and necessary conditions for $\{\tau^*(\cdot)\}$ to be progressive. Since

$$\tau^{*'}(c^*(\theta)) = -\frac{d}{dc^*(\theta)}\frac{c^*(\theta)}{\mathbb{E}[\tilde{y}|l^*(\theta)]} = \frac{c^*(\theta)p'(l^*(\theta))(H-L)\frac{dl^*(\theta)}{dc^*(\theta)}-\mathbb{E}[\tilde{y}|l^*(\theta)]}{\mathbb{E}^2[\tilde{y}|l^*(\theta)]}$$

and

$$\frac{dl^*(\theta)}{dc^*(\theta)} = \frac{\frac{dl^*(\theta)}{d\theta}}{\frac{dc^*(\theta)}{d\theta}} = \frac{\frac{dl^*(\theta)}{d\theta}}{\frac{1}{u'(c^*(\theta))}[v_l(l^*(\theta),\theta)\frac{dl^*(\theta)}{d\theta}+v_\theta(l^*(\theta),\theta)-v_\theta(l^*(\theta),\theta)]} = \frac{u'(c^*(\theta))}{v_l(l^*(\theta),\theta)}$$

we conclude that $\tau^{*'}(\cdot) > 0$ if and only if

$$\frac{\mathbb{E}[\tilde{y}|l^*(\theta)]}{c^*(\theta)} < p'(l^*(\theta))(H-L)\frac{dl^*(\theta)}{dc^*(\theta)} = \frac{p'(l^*(\theta))u'(c^*(\theta))}{v_l(l^*(\theta,\theta))}(H-L)$$

□

References

Albanesi, S., and C. Sleet (2006): Dynamic Optimal Taxation with Private Information, *Review of Econoimc Studies*, 73(1), 1-30.

Atkeson, A., and R., Lucas (1992): On Efficient Distribution with Private Information, *Review of Econoimc Studies*, 59(3), 427-453.

Azariadis, C. (1975): Implicit Contracts and Underemployment Equilibria, *Journal of Political Economy*, 83(6), 1183-1202.

Baily, M. (1974): Wages and Employment under Uncertain Demand, *Review of Economic Studies*, 41(1), 37-50.

Battaglini, M., and S. Coate (2008): Pareto Efficient Income Taxation with Stochastic Abilities, *Journal of Public Finance*, 92(2), 844-868.

Boadway, R., M. Leite-Monteiro, M. Marchand and P. Pestieau (2006): Social Insurance and Redistribution with Moral Hazard and Adverse Selection, *Scandinavian Journal of Economics*, 108(2), 279-298.

Chetty, R., and E. Saez (2010): Optimal Taxation and Social Insurance with Endogenous Private Insurance, *American Economic Journal: Economic Policy*, 2(2), 85-114.

da Costa, C., and I. Werning (2008): On the Optimality of the Friedman Rule with Heterogeneous Agents and Nonlinear Income Taxation, *Journal of Political Economy*, 116(1), 82-112.

Farhi, E., and I. Werning (2010): Progressive Estate Taxation, *Quarterly Journal of Economics*, 125(2), 635-673.

Froot, K.A., D. S. Scharfstein and J. C. Stein (1993): Risk Management: Coordinating Corporate Investment and Financing Policies, *Journal of Finance*, 48(5), 1629-1658.

Golosov, M., N. Kocherlakota and A. Tsyvinski (2003): Optimal Indirect and Capital Taxation, *Review of Economic Studies*, 70(3), 569-587.

Golosov, M., and A. Tsyvinski (2006): Designing Optimal Disability Insurance: A Case for Asset Testing, *Journal of Political Economy*, 114(2), 257-269.

——— (2007): Optimal Taxation with Endogenous Insurance Markets, *Quarterly Journal of Economics*, 122(2), 487-534.

Golosov, M., A. Tsyvinski and I. Werning (2007): New Dynamic Public Finance: A User's Guide, in *NBER Macroeconomics Annual*, ed. by D. Acemoglu, K. Rogoff and M. Woodford. Cambridge, MA: MIT Press, Volume 21, 317-388.

Green, E. (1987): Lending and the Smoothing of Uninsurable Income, in *Contractual Arrangements for Intertemporal Trade*, ed. by E. Prescott and N. Wallace, Minneaopolis: University of Minnesota Press.

Grochulski, B., and N. Kocherlakota (2010): Nonseparable Preferences and Optimal Social Security Systems, *Journal of Economic Theory*, 145(6), 2055-2077.

GUISO, L., L. PISTAFERRI AND F. SCHIVARDI (2005): Insurance within the Firm, *Journal of Political Economy*, 113(5), 1054-1087.

JACKSON, M. O., AND H. SONNENSCHEIN (2007): Overcoming Incentive Constraints by Linking Decisions, *Econometrica*, 75(1), 241-257.

KOCHERLAKOTA, N. (2005): Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation, *Econometrica*, 73(5), 1587-1621.

——— (2010): The New Dynamic Public Finance, Princeton, NJ: Princeton University Press.

LI, A. (2013): Contracting with Disagreement on Deserved Performance Evaluation and Compensation, *Working Paper*, Washington University in St. Louis.

MCDIARMID, C. (1989): On the Methods of Bounded Differences, in Surveys in Combinatorics, ed. by J. Siemons, London Mathematical Society Lecture Note Series 141, Cambridge: Cambridge University Press.

MILGROM, P., AND I. SEGAL (2002): Envelope Theorems for Arbitrary Choice Sets, *Econometrica*, 70(2), 583-601.

MIRRLEES, J. (1971): An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies*, 38(2), 175-208.

NETZERA, N., AND F. SCHEUER (2007): Taxation, Insurance, and Precautionary Labor, *Journal of Public Economics*, 91(7-8), 1519-1531.

RADNER, R. (1981): Monitoring Cooperative Agreements in a Repeated Principal-Agent Relationship, *Econometrica*, 49, 1127-1148.

Rothschild, R., and J. Stiglitz (1976): Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information, *Quarterly Journal of Economics*, 90(4), 629-649.

Rubinstein, A., and M. Yaari (1983): Repeated Insurance Contracts and Moral Hazard, *Journal of Economic Theory*, 30(1), 74-97.

Prescott, E., and R. Townsend (1984): Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard, *Econometrica*, 52(1), 21-46.

Sun, Y. (2006): The Exact Law of Large Numbers via Fubini Extension and Characterization of Insurable Risks, *Journal of Economic Theory*, 126(1), 31-69.

Townsend, R. (1982): Optimal Multiperiod Contracts and the Gain from Enduring Relationships under Private Information, *Journal of Political Economy*, 90(6), 1166-1186.