Optimal Taxation with Permanent Employment Contracts

Paweł Doligalski*

October 12, 2015

Latest version:
http://pdoligalski.github.io/files/optimal_taxation_permanent_contracts.pdf

Abstract

I introduce permanent employment contracts into the dynamic taxation framework. The optimal tax system becomes remarkably simple, as the government outsources most of the insurance provision to employers and focuses mainly on redistribution. Furthermore, I describe a novel trade-off between insurance and redistribution: the permanent contracts facilitate insurance provision, but hamper redistribution to the poor. When the government is progressive, a dual labor market may be optimal. Less productive workers use the fixed-term contracts and are partially insured by the government. On the other hand, the more productive types enjoy full insurance of the permanent contracts, but pay much higher taxes.

1 Introduction

People's lifetime incomes differ both due to initial heterogeneity in earning potential and luck occurring during the working life.[1] These two sources of inequality imply two different roles for income taxation. First, taxes should redistribute income among initially different people according to some welfare criterion. Second, the tax system should provide insurance against differential luck realizations among people that were initially similar. The modern taxation theory, called the New Dynamic Public Finance (NDPF), arrived at the conclusion that the public provision of insurance is difficult due to limited information available to the government. This information problem leads to very complicated, history dependent tax systems.[2] On the other hand, the theory neglects the role of employing firms both with respect to insurance provision and income-shifting responses to taxes.[3] In this paper I show that consideration of both public insurance via income tax and private insurance provided by employers leads to the optimal tax system which is remarkably simple. Furthermore, I describe a novel trade-off between redistribution and insurance. Although the private insurance

*European University Institute. Contact: pawel.doligalski[at]eui.eu.

[1] Huggett et al. [2011] estimate that out of the two, initial differences account to more than 60% of the inequality in lifetime earnings.

[2] NDPF was surveyed by Golosov et al. [2007] and Kocherlakota [2010].

[3] This point was made forcefully by Kopczuk and Slemrod [2006].

markets reduce the random variation of income along the life-cycle, they hamper the government's ability to redistribute. A progressive government may wish to strip less productive workers of the efficient private insurance in order to increase the scale of redistribution.

I study the optimal taxation problem of an economy in which the private sector actively participates in insurance provision. In my model economy risk neutral firms compete for risk averse workers. Workers' productivity is idiosyncratic, evolves stochastically over time and is observed by all firms. Furthermore, workers can save in a risk-free asset and borrow up to a borrowing limit. I follow the Mirrleesian approach in assuming that the government is limited by available information. It implies that the tax system is based on observables: labor income and asset holdings. On the other hand, I postulate that the frictions faced by market participants are of different nature. Private sector actors naturally have better information than the government, but what constrains them is inability to commit. Workers are free by law to change employers. Firms can, at some cost, fire employees. I study the optimal taxation problem under different assumptions on commitment.

If neither workers nor employers can commit, any provision of insurance would lead to a termination of relationship. Insuring a worker essentially means that the firm pays him more than his output in bad times and less than his output in good times. If both a firm and a worker can terminate the relationship anytime at no cost and start a new one with a clean slate, no insurance is possible. In this case the labor market works as a sequence of spot labor markets, the environment studied by the New Dynamic Public Finance (NDPF). Optimally, the government steps in with taxation that both redistributes and insures. However, as the state is constrained by available information, it has to employ a complicated, history dependent income tax system to screen evolving productivities of workers. Although agents are heterogenous only in labor productivity, the optimal insurance provision with private information requires levying a tax on savings (Golosov et al. [2003], Kocherlakota [2005]).

On the opposite extreme is a full commitment economy, in which both workers and firms can credibly promise not to terminate the relationship in the future. In the optimum, the government outsources the entire provision of insurance to employers. The only remaining role of taxes is redistribution. I show that the optimal redistributive taxes in this dynamic environment are given by the optimal 'static' tax formulas of Mirrlees [1971] model, derived by Saez [2001]. Hence, I provide microfundations for interpreting the static Mirrlees [1971] model as a taxation of lifetime income. Although the elasticities relevant in the tax formula are the appropriate lifetime averages, only the initial distribution of types matters for taxation. Intuitively, if all people enter the labor market with an identical initial productivity and the same distribution of future shocks, any inequality in income would be a matter of insurance, not redistribution.

Probably the most empirically relevant case to consider is the one-sided commitment, in which firms can commit not to fire, but workers cannot commit to stay.[4] Harris and Holmstrom [1982] showed that when workers cannot borrow, the labor compensation is downwards rigid and increases anytime a worker is tempted to quit.[5] They also note that insurance provision increases with

[4] See for instance Chiappori et al. [1999].

[5] They assume that a worker's productivity is fixed yet unknown, while worker and firms symmetrically learn it over time. Since the optimal contract depends on the posterior mean of productivity, which is a random variable, this framework is equivalent to mine in which the productivity is observed, but evolves stochastically.

worker's ability to borrow. Borrowing allows firms to offer more backloaded compensation, which gives workers more incentives to stay with the current employer. With sufficient borrowing limit the firms can offer workers full insurance, just as in the full commitment case. Furthermore, I show that even if workers cannot borrow, the government can implement full consumption insurance by properly adjusting the taxes. Essentially, the tax payments have to substitute for the missing trade in the risk-free bond. Hence, under one-sided commitment the government can implement any full commitment allocation, although the implied structure of labor income may be very different. It will have important consequences for decentralization of the optimal allocation.

In order to link the environments studied above to the labor regulation, I consider a framework in which the firms' commitment depends on the firing cost specified in the contract. Permanent labor contracts with high firing cost allow firms to commit not to fire workers. On the other hand, fixed-term contract with no firing cost imply no commitment of firms. By differential taxation of labor contract with different firing cost, the government can control the commitment friction faced by each type of worker. Naturally, when the government is concerned predominantly with insurance provision, it is optimal to equip all firms with ability to commit by promoting the permanent contracts. However, when the government has a strong will to redistribute to the initially less productive, these workers may optimally end up with fixed-term contracts and no firm insurance. The reason behind this finding is following. When the employer can commit, she structures worker's income in order to minimize the worker's tax burden. The currently productive workers would benefit from shifting income to the future and claiming transfers due to low current earnings. Such income shifting is impossible without the firm's commitment. Hence, when low earners are employed on the fixed-term contract with no firm's commitment, the government is able to tax the high earners more.

How to decentralize the optimal allocation? First, the redistributive tax of permanent workers should not be based on labor income or total income which can exhibit some time variation, either due to limited commitment friction or due to trends in productivity. Instead, the tax should be based instead on consumptions expenditures which remain stable through time. When the tax is progressive, it can be based only on current consumption expenditures - no history dependence is required. Savings of the permanent workers should not be taxed, unless it is spend on current consumption. On the other hand, the fixed-term workers receive no private insurance and are covered by a public insurance program. As in NDFP, these programs involve taxation of savings, which I interpret is as means tests that guarantee that only asset poor workers are eligible for income support. Such an arrangement, in which the majority of workers face a redistributive 'static' tax, while the fixed-term workers who tend to have low income use social insurance programs, closely resembles the tax and transfer systems in most developed countries.

The endogenous commitment framework yields testable implications regarding the income risk of workers employed on different labor contracts. I verify these implications with the administrative dataset of employment histories from Italy. I show that the residual income variance is higher by 78% for fixed-term workers in comparison to the permanent workers. This estimate is consistent with the theory and economically significant. There are many studies documenting the impact of fixed-term contracts on the employment risk (see literature review in Kosior et al. [2015]). To the best of my knowledge, I am the first to document the impact of fixed-term contracts on income risk, conditional on staying employed. A proper causal analysis of the link between firing costs and insurance provided by firm is an interesting topic for future research.

Finally, I calibrate a simple life-cycle model to Italy. I find that the utilitarian government would ban the fixed-term contracts. On the other hand, the Rawlsian government that cares only about the least well-off workers would introduce a dual labor market: the initially less productive workers are employed on the fixed-term basis, while the more productive workers are employed permanently. I compare social welfare of the constrained efficient allocation with possible permanent contracts to the best allocation when only fixed-term contract are allowed. Permanent contracts allow the government to increase welfare gains from taxation by one-half (utilitarian case) or one-third (Rawlsian case).

Related literature. New Dynamic Public Finance (NDPF), also called the dynamic Mirrleesian literature, studies an optimal taxation problem in a dynamic economy with private information. Golosov et al. [2003] and Kocherlakota [2005] show that in this framework capital taxation plays an important role, as it deters dynamic incentives for lying about one's productivity. NDPF is criticized for ignoring private insurance markets and prescribing very complicated, history dependent tax systems. I show that by including endogenous private markets the optimal tax system becomes simple and consistent with the original static Mirrlees [1971] model. Besides the optimal tax on total income net of savings, there is no need for an additional capital taxation, since agents reveal their type only once, in the initial period.

There are two related papers that study the equilibrium insurance provision by firms. Golosov and Tsyvinski [2007] consider an environment where firms do not observe workers' productivity. Moreover, firms and workers can fully commit and sign contract 'behind a veil of ignorance', before workers draw their first productivity realization. They show that employers can replace government in both insuring and redistributing workers, while the government can improve the allocation by correcting excessive capital accumulation. Ales and Maziero [2009] also assume that firms do not observe worker's productivity. Moreover, in their framework contracts are non-exclusive - an employer cannot prevent employees from having a side deal with another firm. Non-exclusivity makes screening of workers impossible and equilibrium is equivalent to a sequence of spot labor markets with possibility of trading in a risk-free asset. Both papers assume that firms and workers can fully commit to labor contracts. Instead, I stress the importance of the lack of commitment, a realistic feature of market transactions. It can account for private lack of insurance, even if employers observe workers' productivity.

Harris and Holmstrom [1982] propose the setting in which a worker and a firm symmetrically learn about a worker's productivity. Firm is able to commit to a contract, while the worker will always leave the firm. Under assumption of no of workers' borrowing, the equilibrium wage is rigid: it does not fall and it raises only when the worker gets a competing offer from the other firm. When workers can borrow sufficiently much, they are fully insured by firms. In the section on one-sided commitment I find the constrained efficient allocation in such an economy with government that does not observe the posterior mean of each worker's productivity. Other limited commitment models of labor market were surveyed by Thomas and Worrall [2007]. The one-sided commitment is important also in other insurance markets. Hendel and Lizzeri [2003] show that the data on life insurance market is consistent with one-sided limited commitment of consumers: premiums are front-loaded and more front-loaded contracts coincide with lower rate of consumers switching contracts. Handel et al. [2013] use similar framework to model a health exchange.

Structure of the paper. The next section introduces the environment and sets up the general government's taxation problem. The sections that follow describe the optimal allocation in the cases without commitment, with full commitment, with one-sided commitment and with one-sided endogenous commitment. In Section 7 I validate the predictions of the model with the administrative dataset from Italy. Then I calibrate the model to Italy and describe the optimal taxation and labor regulation in Section 8. The last section concludes. All the proofs are placed in the Appendix A.

2 Framework

2.1 Firms and workers

There is a continuum of workers. Workers live for $\mathcal{T} \in \mathbb{N}_+$ periods. In each period they draw a productivity, which I describe in detail below. A worker with productivity θ and labor supply n produces output θn. Workers sell their labor to firms in exchange for a labor income y. Workers have access to the risk-free asset, in which they can save and borrow up to the limit $b \geq 0$ at the gross interest rate R. I denote worker's choice of assets by a and assume that workers have no wealth initially. A worker's contemporaneous utility depends on consumption and labor supply according to a twice differentiable function $U(c, n) = u(c) - v(n)$, where u is increasing and strictly concave, while v is increasing and strictly convex. Worker's lifetime utility is a discounted expected stream of contemporaneous utilities, where β is a discount factor. I assume that the interest rate is consistent with worker's discounting: $R = \beta^{-1}$.

There is a continuum of identical firms. Firms maximize expected profits by hiring workers, compensating them with labor income and collecting output. Firms observe each worker's productivity and labor supply. I assume no entry cost for firms.

2.2 Productivity histories

In each period period t (where $1 \leq t \leq \mathcal{T}$) a worker draws productivity from a finite set $\Theta_t \subset \mathbb{R}_+$. A history is a tuple of consecutive productivity draws starting at the initial period. I denote by $|h|$ the length of the history h, i.e. the number of productivity draws it contains. The history h belongs to the set $\Theta^{|h|} = \Pi_{t=1}^{|h|} \Theta_t$ and the set of all histories is $\Theta = \cup_{t=1}^{\mathcal{T}} \Theta^t$. Since all histories start in period 1, the length of the history is also the current time period. The i-th element of history h is h_i and the tuple of its first i elements is $h^i = (h_1, ..., h_i)$. In order to simplify notation, I denote the last productivity at the history h as $\theta(h) \equiv h_{|h|}$ and the history directly preceding the history h as $h^{-1} \equiv h^{|h|-1}$. For clarity, consider a following example:

$$h = (\theta_a, \theta_b, \theta_c) \in \Theta^3,\ |h| = 3,\ h^{-1} = (\theta_a, \theta_b),\ \theta(h) = \theta_c.$$

The probability of drawing some history h of length t is equal $\mu(h)$ which is non-negative and sums up to 1 for all histories of this length: $\forall_t \sum_{s\in\Theta^t} \mu(s) = 1$. The cumulative distribution function is defined as $M(h) \equiv \sum_{\theta\in\Theta_{|h|}\cap[0,\theta(h)]} \mu\left(h^{-1}, \theta \mid h^{-1}\right)$. In practice, I will work mostly on the collections of histories of positive probability, denoted by $\mathcal{H} \equiv \{h \in \Theta : \mu(h) > 0\}$. $\mathcal{H}_t$ is the set of histories with positive probability of length t. By $\mathcal{X}(h)$, where $\mathcal{X}$ is a set of histories and $h \in \mathcal{H}$, I mean the subset of elements of $\mathcal{X}$ that contain $h : \mathcal{X}(h) = \left\{s \in \mathcal{X} : s^{|h|} = h\right\}$. Specifically, $\mathcal{H}_t(h)$ is the set of possible histories of length t that contain sub-history h. The probability of drawing history $s \in \mathcal{H}(h)$ conditional on history h, where $\mu(h) > 0$, is

$$\mu(s \mid h) = \frac{\mu(s)}{\sum_{s'\in\mathcal{H}_{|s|}(h)} \mu(s')}.$$

2.3 Allocation

Definition 1. The allocation (y, a, T, n) specifies

- the labor income $y : \mathcal{H} \to \mathbb{R}$,
- assets $a : \mathcal{H} \to [-b, \infty)$, which cannot be negative in the last period: $\forall_{h\in\mathcal{H}_T} a(h) \geq 0$,
- total tax $T : \mathcal{H} \to \mathbb{R}$,
- labor supply $n : \mathcal{H} \to \mathbb{R}_+$.

The expected utility of a worker at the history history $h \in \mathcal{H}$, given the allocation (y, a, T, n) is

$$\mathbb{E}\mathbf{U}_h(y, a, T, n) \equiv \sum_{s\in\mathcal{H}(h)} \mu(s \mid h)\, \beta^{|s|-|h|} U(c(h), n(s)), \tag{1}$$

where the consumption at history $h \in \mathcal{H}$ is determined by the usual budget constraint

$$\forall_{h\in\mathcal{H}\setminus\mathcal{H}_1}\ c(h) = y(h) + Ra\left(h^{-1}\right) - a(h) - T(h).$$

I assume that workers initially have no wealth, so consumption at the initial histories is

$$\forall_{h\in\mathcal{H}_1}\ c(h) = y(h) - a(h) - T(h).$$

By $\mathbb{E}\mathbf{U}(y, a, T, n)$ I denote the expected utility from the *ex ante* perspective

$$\mathbb{E}\mathbf{U}(y, a, T, n) \equiv \sum_{h\in\mathcal{H}_1} \mu(h)\, \mathbb{E}\mathbf{U}_h(y, a, T, n).$$

Profits from hiring a worker at the history h given the allocation (y, a, T, n) are

$$\mathbb{E}\bar{\pi}_h(y, n) = \sum_{s \in \mathcal{H}(h)} \mu(s|h) R^{|h|-|s|} (\theta(s) n(s) - y(s)). \tag{2}$$

2.4 Direct mechanism

I assume that the government observes labor income y and asset trades a', but does not observe the productivity θ, hours worked n or individual output θn. In order to write and solve the optimal tax problem, I use the revelation principle, according to which without the loss of generality I can focus on direct mechanisms. The government set up a direct mechanism which in each period collects type reports of workers and assigns them labor income y, assets a and taxes T. As the government do not observe individual labor supply, it is determined in the equilibrium corresponding to the governments choice (y, a, T), which I describe below.

The reporting strategy of the worker is denoted by r, where $r : \mathcal{H} \to \mathcal{H}$. I impose the consistency condition that $\forall_{s,h \in \mathcal{H}} s \in \mathcal{H}(h) \implies r(s) \in \mathcal{H}(r(h))$. It simply means that consecutive history reports cannot be at odds with which histories are in fact possible. Let's denote the set of consistent reporting strategies by $\mathcal{R}$. The truthful reporting strategy r^* is $r^*(h) = h$ for all $h \in \mathcal{H}$.

Let's define the expected utility of a worker at the history h, given the allocation (y, a, T, n) and the reporting strategy r

$$\mathbb{E}U_h(y, a, T, n, r) \equiv \mathbb{E}\mathbf{U}_h(y \circ r, a \circ r, T \circ r, n) \tag{3}$$

where $f \circ r$, for some function $f : \mathcal{H} \to \mathbb{R}$, is a composite function: $\forall_{h \in \mathcal{H}} (f \circ r)(h) = f(r(h))$. In words, the worker's payoff from a reporting strategy r at allocation (y, a, T, n) is given by an expected utility from the allocation $(y \circ r, a \circ r, T \circ r, n)$. Note that r affects (y, a, T) which are assigned by the government. It does not affect labor supply n, as it is decided in equilibrium by firms that observe worker's type.

Let's similarly define firm's profits from allocation (y, a, T, n) at the reporting strategy

$$\mathbb{E}\pi_h(y, n, r) \equiv \mathbb{E}\bar{\pi}_h(y \circ r, n).$$

2.5 Equilibrium

Definition 2. A labor contract is a duple (n, r).

Firms typically compete by offering contract specifying labor supply and labor income.[6] Since I use the direct mechanism approach and the labor income is observed by the government, the firm cannot offer any structure of labor income. Instead, firms offer reporting strategies which, together with y specified by the government, constitute a map from type to a labor income. Given y, firms

[6] See Appendix **??** for description of equilibrium without the government.

essentially compete by offering the effective labor income $y \circ r$. Such formulation highlights the importance of firms in transmitting the information to the government, emphasized by Kopczuk and Slemrod [2006].

Definition 3. A labor contract is self-enforcing if at no history the worker or the firm have incentives to terminate the contract.

The notion of self-enforcing contract will be defined more clearly in each case I consider.

Definition 4. A labor contract (n,r) is an equilibrium contract given (y,a,T) if (i) (n,r) yields non-negative profits: $\mathbb{E}\pi_h(y,n,r) \geq 0$, (ii) (n,r) is self-enforcing, (iii) given (n,r) no other self-enforcing labor contract yields positive profits. The set of equilibrium labor contracts given (y,a,T) is $\mathcal{C}(y,a,T)$.

2.6 The general government's problem

The government chooses the mechanism in order to maximize the social welfare function

$$\max_{(y,a,T)} \sum_{\theta\in\mathcal{H}_1} \lambda(\theta)\,\mu(\theta)\,\mathbb{E}U(y,a,T,n,r^*), \tag{4}$$

where λ are the Pareto weights with the average value equal to 1: $\sum_{\theta\in\mathcal{H}_1}\lambda(\theta)\,\mu(\theta) = 1$. The optimization is subject to the government budget constraint

$$\sum_{h\in\mathcal{H}} \mu(h)\,R^{-|h|}T(h) \geq 0 \tag{5}$$

and the equilibrium constraint, which means that (n,r^*) is the equilibrium contract, given the government's choice

$$(n,r^*) \in \mathcal{C}(y,a,T). \tag{6}$$

3 No commitment

If there is no commitment, at no point neither firm nor worker can potentially gain from terminating the labor contract. It implies that expected profits of the firm have to be equal to zero at each history.

Lemma 1. *Under no commitment, the set of equilibrium contracts is*

$$\mathcal{C}^{NC}(y,a,T) \equiv \arg\max_{r,n} \mathbb{E}U(y,a,T,n,r), \ \ s.t.\ \forall_{h\in\mathcal{H}}\mathbb{E}\pi_h(y,n,r) = 0.$$

In any equilibrium contract $(r,n) \in \mathcal{C}^{NC}(y,a,T)$ at any history $h\in\mathcal{H}$ the worker's labor income is equal to the worker's output.

$$y(r(h)) = \theta(h)\,n(h).$$

Lemma 1 tells us that the reporting strategy uniquely determines the equilibrium labor supply policy, since output equals labor supply in each period. Hence, we can reformulate the equilibrium constraint (6) as

$$\forall_{r\in\mathcal{R}}\mathbb{E}U\left(y,a,T,r^{*},n\left(r^{*}\right)\right)\geq\mathbb{E}U\left(y,a,T,r,n\left(r\right)\right),\ \text{where }\forall_{h\in\mathcal{H}}n\left(r\left(h\right)\right)=\frac{y\left(r\left(h\right)\right)}{\theta\left(h\right)}. \tag{7}$$

Corollary 1. *Under no commitment, the government's problem is a New Dynamic Public Finance taxation problem.*

No commitment prevents firms from providing any insurance, hence in the optimum the government uses taxes to both redistribute and insure. Since the government does not observe worker's productivity, provision of insurance becomes a dynamic screening problem. The resulting optimum features imperfect consumption insurance and taxes that depend on the whole history of reports in a complicated way. The labor supply is typically distorted downwards. Furthermore, the government taxes savings of people with reported low productivity. It discourages the following double deviation: produce and save a lot if the current productivity is high and in the future report low productivity in any contingency.

4 Full commitment

Under full commitment, the firm is offering contract (r,n) in order to maximize utility of some initial type $\theta_1\in\times_1$, subject to non-negative profits. Since both workers and firms commit, we do not have to worry about incentives to terminate the relationship.

Lemma 2. *Under full commitment, the set of equilibrium contracts is*

$$\mathcal{C}^{FC}\left(y,a,T\right)\equiv\arg\max_{r,n}\mathbb{E}U\left(y,a,T,n,r\right),\ \ s.t.\ \forall_{h\in\mathcal{H}^{1}}\mathbb{E}\pi_{h}=0.$$

In contrast to Golosov and Tsyvinski [2007], I preclude contracting before the initial type is realized. As a result, the firms do not redistribute income among the initial types. It leaves room for the redistributive taxation.

The firm will offer a reporting strategy which is the most beneficial to type θ_1 and corresponding labor supply policy which minimizes the disutility cost of working

$$\forall_{h\in\mathcal{H}}\frac{v'\left(n\left(h\right)\right)}{\theta\left(h\right)}=\frac{v'\left(n\left(h_{1}\right)\right)}{\theta\left(h_{1}\right)}. \tag{8}$$

The zero-profit condition means that the expected lifetime output of any initial type $\theta_1\in\Theta_1$ is equal to the expected lifetime income

$$\sum_{h\in\mathcal{H}(\theta_{1})}R^{1-|h|}\mu\left(h\right)\theta\left(h\right)n\left(h\right)=\sum_{h\in\mathcal{H}(\theta_{1})}R^{1-|h|}\mu\left(h\right)y\left(r\left(h\right)\right). \tag{9}$$

Note that output has to equal income only in the expectations over the lifetime. It means that, conditional on consumption allocation being fixed, the worker is indifferent between two allocations of effective income $y \circ r$ as long as they have the same expected lifetime value. I denote by $\mathbb{E}Y(r)$ the expected lifetime income from a particular reporting strategy (the right hand side of (9)).

Theorem 1. *Optimum under full commitment in the labor market is characterized by: (i) full consumption insurance:* $\forall_{h\in\mathcal{H}} c(h) = c(h_1)$*, (ii) lifetime labor income that is a deterministic function of initial type:* $\forall_{h\in\mathcal{H}_{\mathcal{T}}(\theta_1)} \sum_{t=1}^{\mathcal{T}} R^{1-t} y(h^t) = Y(h_1)$.

Let's define the indirect utility function

$$V_\theta(C,Y) = \sum_{t=1}^{\mathcal{T}} \beta^{t-1} u\left(\left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)^{-1} C\right) - \sum_{h\in\mathcal{H}(\theta)} \beta^{|h|-1} v(n(h)),$$

where $n : \mathcal{H}(\theta) \to \mathbb{R}_+$ satisfies (8) and $\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h)\theta(h)n(h) = Y$. By Theorem 1, we can reformulate the government problem in the following way: for each initial type $\theta_1 \in \Theta_1$ the government chooses lifetime consumption $C(\theta_1) = \sum_{t=1}^{\mathcal{T}} R^{1-t} c(\theta_1)$ and lifetime income $Y(\theta_1)$ in order to maximize the social welfare

$$\max_{(C(\theta),Y(\theta))_{\theta\in\Theta_1}} \sum_{\theta\in\Theta_1} \lambda(\theta)\mu(\theta) V_\theta(C(\theta), Y(\theta))$$

subject to the budget constraint

$$\sum_{\theta\in\Theta_1} (Y(\theta_1) - C(\theta_1)) \geq E$$

and the initial incentive-compatibility constraints

$$\forall_{\theta,\theta'\in\Theta_1}\ V_\theta(C(\theta), Y(\theta)) \geq V_\theta(C(\theta'), Y(\theta')),$$

Note that it is just a static taxation problem studied by Mirrlees [1971]. Let's define the lifetime tax as $\tilde{T}(\theta) \equiv Y(\theta) - C(\theta)$. Diamond [1998] and Saez [2001] express the optimal taxes in this environment, under assumption that the preferences satisfy a single-crossing condition. Intuitively, it means that the more productive worker is willing to produce more than the less productive worker under the same circumstances. Lemma 3 provides sufficient conditions for V_θ to satisfy the single-crossing condition.

Assumption 1. *For any length of the history t and for any sequence of productivity draws following the first draw* $s \in \times_{i=2}^{t}\Theta_i$ $M(\theta_1, s) \geq M(\theta_2, s)$ iff $\theta_2 > \theta_1$.

Lemma 3. *Under Assumption 1 the function* $V_\theta(C,Y)$ *satisfies the single-crossing condition.*

Assumption 1 means that the future distribution of productivity of initially higher type weakly first order stochastically dominates the distribution of initially less productive type. In other words, the initial high productivity is not a signal of low productivity in the future, relative to other workers.

In order to apply the existing optimal tax formulas, we need to slightly modify the environment - the initial distribution of types is continuous. Then, under the single-crossing and the monotonicity condition, the optimal marginal tax rates are described by the formula of Saez [2001].

Assumption 2. *Θ_1 is an interval of real numbers. A probability density function over Θ_1 is $\tilde{\mu}(\theta)$ and the cumulative distribution function is $\tilde{M}(\theta)$.*

Theorem 2. *Under Assumptions 1 and 2, if the implied lifetime income schedule $Y(\theta)$ is non-decreasing, the optimal lifetime tax of an initial type $\theta \in \Theta_1$ satisfies*

$$\frac{\tilde{T}'(\theta)}{1-\tilde{T}'(\theta)} = \frac{1-\tilde{M}(\theta)}{\theta\tilde{\mu}(\theta)}\frac{1+\bar{\zeta}^u(\theta)}{\bar{\zeta}^c(\theta)}\mathbb{E}\left\{(1-\omega(\theta'))\, e^{\int_\theta^{\theta'}\frac{\bar{\xi}(\theta'')}{\bar{\zeta}^c(\theta'')}\frac{Y'(\theta'')}{Y(\theta'')}d\theta''}\middle|\, \theta' \geq \theta, \theta' \in \Theta_1\right\}$$

where $\bar{\zeta}^c(\theta) = \sum_{h\in\mathcal{H}(\theta)} R^{1-|h|}\mu(h \mid \theta)\frac{\theta(h)n(h)}{Y(\theta)}\zeta^c(h)$ is the weighted lifetime average of the compensated elasticity of labor supply, $\bar{\xi}(\theta) = \left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)^{-1}\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)\,\xi(h)$ is the lifetime average wealth effect, $\bar{\zeta}^u(\theta) = \bar{\zeta}^c(\theta) + \bar{\xi}(\theta)$ is the lifetime average uncompensated elasticity of labor supply and $\omega(\theta) = \frac{\lambda(\theta)u'(c(\theta))}{\eta}$ is the marginal social welfare weight of the initial type θ.

The optimal marginal tax rates depend on the distribution of types, the labor supply elasticities as well as social preferences. As the government is concerned only with redistribution, the marginal tax rates depend directly only on the initial distribution of types. Intuitively, if each worker had the same initial productivity, there would be no scope for the redistributive taxation - any inequality of income would be matter of insurance. Furthermore, the elasticities that enter the tax formula are the lifetime averages. Specifically, the lifetime compensated elasticity of labor supply is an average elasticity of labor supply, weighted by the current output.

Definition 5. Define the laissez-faire Pareto weights $\lambda^{lf}(\theta)$ as

$$\forall_{\theta,\theta'\in\Theta_1}\frac{\lambda^{lf}(\theta)}{\lambda^{lf}(\theta')} = \frac{u'(c(\theta'))}{u'(c(\theta))}, \quad \int_{\Theta_1}\lambda^{lf}(\theta)\,d\theta = 1,$$

where $c(\theta)$ is one period consumption of the initial type θ in the full commitment equilibrium under no government intervention.

Corollary 2. *The optimum under full commitment Pareto dominates the optimum under no commitment when the Pareto weights are sufficiently close to λ^{lf}.*

With the laissez-fare Pareto weights the government does not want to redistribute between initial types. However, the government still cares about insurance. Since full commitment allows for the full insurance, while under no commitment in the optimum insurance is imperfect, the optimum in the former case Pareto dominates the latter.

5 One-sided commitment

In this section I describe the taxation problem, where firms can commit, but workers cannot. It seems to be the most empirically relevant case.[7] As the workers cannot commit, it cannot be the case that they ever have incentives to leave the firm. It is equivalent to a requirement that expected profits of firms are always non-positive.

Lemma 4. *Under one-sided commitment of firms, the set of equilibrium contracts is*

$$\mathcal{C}^{OC}(y,a,T) \equiv \arg\max_{r,n} \mathbb{E}U(y,a,T,n,r), \ \ s.t. \ \forall_{\theta_1\in\Theta_1}\mathbb{E}\pi_{\theta_1} = 0 \ and \ \forall_{h\in\mathcal{H}}\mathbb{E}\pi_h \leq 0.$$

As Harris and Holmstrom [1982] point out, firms can provide full insurance under the one-sided commitment if workers can borrow sufficiently much. Hence, once borrowing is sufficient, the government will be able to implement the full-commitment allocation with a 'static' Mirrleesian tax. Importantly, this tax should be based not on labor income, but on consumption expenditures. In a static model both taxes are equivalent, but in a dynamic model we need to consider savings and borrowing as well. As borrowing plays important role in provision of insurance, it shouldn't be deterred by the tax system and should be taxes only at the point where it is translated into consumption.

When borrowing is insufficient, the equilibrium without the government intervention will not involve full insurance. However, as the asset traded is not contingent on individual type, the government can substitute missing trading opportunity with age-dependent taxation. In this way the tax smooths the consumption financed by the backloaded income stream.

Proposition 1. *Under one-sided commitment, the government can implement any allocation under full commitment.*

When the structure of the model is sufficiently simple, permanent contracts Pareto dominate fixed-term contracts also when social preferences are skewed towards productive types.

Example 1. There are two time periods ($\mathcal{T} = 2$) and two initial productivity levels $\left(\mathcal{H}_1 = \left\{\underline{\theta}, \overline{\theta}\right\}, \ \overline{\theta} > \underline{\theta}\right)$. The productivity distribution in the second period is independent of the first productivity draw.

Proposition 2. *Consider Example 1 and the reporting strategy $r_{\overline{\theta},\underline{\theta}} \in \mathcal{R}$ of initial high type mimicking initial low type ($r_{\overline{\theta},\underline{\theta}}\left(\overline{\theta}\right) = \underline{\theta}$ and $r_{\overline{\theta},\underline{\theta}}(h) = r^*(h) \ \forall_{h\in\mathcal{H}\setminus\overline{\theta}}$). Suppose that in the optimum with no commitment the incentive constraint with respect to strategy $r_{\overline{\theta},\underline{\theta}}$ does not bind. The government can make both initial types better off under one-sided commitment.*

The proof of this proposition relies on the observation that when the government does not introduce downward labor distortions in the first period in the optimum under no commitment, it is possible to replicate exactly the same allocation under one-sided commitment. However, that's not the case when there are downward labor distortions in the first period.

[7] The opposite case, in which firms cannot commit and workers can is in this setting, where firms face no borrowing constraints, less interesting. The equilibrium without the government intervention implements the same allocation as under full commitment. Essentially, the firm can commit not to fire worker by giving him or her large deposit in the initial period. The worker pays back the deposit in the last period iff the firm did not terminate the contract.

I will illustrate the difficulty in implementing downwards distortions under one-sided commitment. Consider the optimum under one-sided commitment in Example 1, where Pareto weight are chosen such that there is no redistribution. In this case the labor supply of any initial type will be undistorted, while the labor supply in the second period will be distorted downwards when productivity is low. Suppose that the government wants to decrease labor supply of the initial type $\underline{\theta}$ by δ. Under no commitment it is enough to decrease $y(\underline{\theta})$ by $\underline{\theta}\delta$. Since current labor is bound to the current income, the labor supply will decrease accordingly. Under one-sided commitment the equilibrium contract minimize the utility cost of production subject to the the non-positive profit constraints in the second period. However, in this case these constraints will not bind and the firm is going to offer the full commitment contract. In the full commitment contract the labor supply is chosen in order to equalize disutility from labor across all states and time. Hence, decreasing $y(\underline{\theta})$ will lead to a decrease of labor supply at all histories.

6 Endogenous commitment

I make the firm's commitment endogenous by introducing a firing cost f, which can be interpreted as a severance pay to the fired worker. Suppose that the legal framework is such that the firms can hire workers on one of the two contracts. With the fixed-term contract the firm may fire worker in each period without any cost: $f = 0$. Essentially, the contract is valid for one period only, after which the firm and the worker are free, but not obliged, to sign another contract at the new terms. With a permanent contract the firing cost is high: $f = \bar{f}$. Specifically, I assume that $\bar{f}$ is so high such that the firm never have incentives to fire workers. Moreover, I assume that the government is able to tax the two contracts differentially. As a result, the government chooses (y, a, T, f) where

$$f : \mathcal{H} \to \{0, \bar{f}\}.$$

Without the loss of generality we can impose that f satisfies

$$f(h) = \bar{f} \implies \forall_{s \in \mathcal{H}(h)} f(s) = \bar{f}.$$

It means that once the worker is hired at the permanent contract, he stays at this contract for the rest of his life. Instead, a fixed-term contract worker can in the next period be hired at the permanent contract or at the fixed-term contract.

Lemma 5. *The equilibrium set of contracts under firing cost f is given by*

$$\mathcal{C}^{EC}(y, a, T, f) \equiv \arg\max_{r,n} \mathbb{E}U(y, a, T, n, r), \ \text{s.t.} \ \forall_{\theta_1 \in \Theta_1} \mathbb{E}\pi_{\theta_1} = 0 \ \text{and} \ \forall_{h \in \mathcal{H}} - f(r(h)) \leq \mathbb{E}\pi_h \leq 0.$$

Note that under endogenous commitment the government can replicate one-sided commitment or no commitment allocations by simply setting the firing cost high or low respectively. The endogenous commitment also allows the government to introduce dual labor market, in which different type of workers are employed under different labor regulation. Dual labor markets exist in many European countries (see e.g. Kosior et al. [2015]).

Proposition 3. *In the optimum some initial types are employed at the permanent contracts.*

As the permanent contracts allow the firms to provide full insurance to workers, they will be always used in the optimum. Perhaps surprisingly, the fixed-term contracts are sometimes optimal as well. In the proposition below for simplicity I consider risk-neutral workers. In this way I abstract from the insurance motive for taxation and focus on redistribution. Note that by continuity the same results holds for a moderate amount of risk-aversion.[8]

Proposition 4. *Consider Example 1. Suppose that workers are risk-neutral: $U(c,n) = c - v(n)$, and the government is Rawlsian: $\lambda(\underline{\theta}) = 1$ and $\lambda(\overline{\theta}) = 0$. There exists a threshold $\tilde{\underline{\theta}} > 0$ such that when $\underline{\theta} < \tilde{\underline{\theta}}$, the government strictly prefers assigning the fixed-term contract to type $\underline{\theta}$.*

Although the permanent contracts provide superior insurance, they may hamper the government's ability to redistribute. As in the setting of Proposition 4, the permanent contracts of workers with initially low productivity make is easy for the high productivity type to mimic them. Under permanent contract the high productivity type may produce substantial amount of output while pretending to be of low productivity, since the firm shift worker's income to the next period. Instead, under the fixed-term contract the firm cannot commit to pay back the output to the worker, hence the worker cannot use the superior productivity while pretending to be the less able type. Under the permanent contracts mimicking becomes easier and the government's screening problem more difficult, which limits the opportunity for redistribution. Introduction of the fixed-term contract relaxes the binding incentive constraint and allows to increase the transfers to the initially low type.

Redistribution is not the only reason to use the fixed-term contracts in the optimum. The fixed-term contracts may be preferable also because they reduce the optimal magnitude of labor distortions. The labor distortions of some type θ are used to deter other type from mimicking θ. If θ receives the fixed-term contract and other workers find it strictly less attractive to mimic him, the government will optimally reduce distortions of this type in comparison to the case of permanent contract.

Proposition 5. *Consider Example 1. Suppose that under permanent contracts the initial type $\underline{\theta}$ optimally does not supply labor: $Y(\underline{\theta}) = 0$. Type $\underline{\theta}$ can obtain strictly greater welfare with the fixed-term contract.*

In Proposition 5 I consider the allocation in which under the permanent contracts the optimal distortions are so severe the initially less productive type does not work at all. It is likely to be the case when the initial productivity difference is substantial and the population share of the initially less productive type is small.

Under the permanent contracts, the government optimally decreases labor income of initially low type even in the second period, since the mimicking high type could produce some output in the first period, using the superior initial productivity. If the initially less productive type receives

[8]For instance, when the utility function is CRRA with a relative risk aversion of σ, then by the continuity Proposition 4 holds also for $0 \leq \sigma < \bar{\sigma} \in \mathbb{R}_{++}$.

the fixed-term contract, such distortions are no longer useful, as the mimicking type cannot use the superior initial productivity. Hence, the government optimally lifts part of labor distortions, increasing welfare of the initially less productive worker.

6.1 Decentralization

In this subsection I show how to decentralize the optimal allocation when all workers receive the permanent contracts with full consumption insurance and hint at what changes when the dual labor markets are optimal. I focus on the case in which the borrowing is sufficiently high, such that the workers can obtain full consumption insurance even in the absence of taxes. Note that the simplest way to decentralize the allocation in which the borrowing is not sufficient is for the government to set up an artificial asset market with the sufficient borrowing limit, where the payment of debt is enforced by the taxation ability of the state.

How to implement allocation with the lifetime consumption $C(\theta)$, lifetime income $Y(\theta)$ and per period tax payment $T(\theta) = \frac{Y(\theta)-C(\theta)}{\sum_{t=1}^{T} R^{1-t}}$ in the simplest way? Define consumption expenditures as $x(h) \equiv y(h) + Ra\left(h^{-1}\right) - a(h)$. As workers receive full consumption insurance via borrowing, the consumption expenditures are constant across time for each initial type: $\forall_h x(h) = x(h_1)$. This property makes it an ideal base for taxation, as in each period a given level of x would correspond to the same level of tax payment T_x.[9] In contrast, other variables such as labor income $y(h)$ or total income $y(h) + Ra(h)$ will exhibit time variation both due to limited commitment friction, which requires a backloading of labor income, or simply because of deterministic time trends in productivity. Implementing a full consumption insurance based for instance on labor income would require the tax payment at each level of y to change in each period, in order to keep the tax burden of a given initial type constant.

Suppose the government wants to implement the allocation (y, a, T). Define the constant level of consumption expenditure tax as $\bar{T}_x = T \circ x^{-1}$, where $x^{-1} : \mathbb{R}_+ \to \Theta_1$ is the inverse function of $x(\theta) \equiv y(\theta) - a(\theta), \theta \in \Theta_1$. One simple way of implementing the tax is to base it only on the initial consumption expenditures only: $T_x\left((x_i)_{i=1}^t\right) = \bar{T}_x(x_1)$, where $(x_i)_{i=1}^t$ is the history of consumption expenditures up to period t. The proposition below shows that when the tax is progressive, it can be made even simpler.

Proposition 6. *Suppose that $\bar{T}_x$ is convex. The allocation can be decentralized with a tax that depends only on current consumption expenditures $T_x\left((x_i)_{i=1}^t\right) \equiv \bar{T}(x_t)$.*

The 'static' tax that depends only on current consumption expenditures is not always feasible, as workers can introduce some volatility in their consumption expenditures. When the tax is regressive (concave), such volatility reduces the average tax burden and agents with sufficiently low risk aversion would be willing to deviate. Note that statutory income tax in virtually any country is (weakly) progressive.

[9] It is recognized in the literature that a non-linear consumption tax is difficult to implement, when the tax depends on a type of a consumption good. However, the tax I propose does not differentiate between different consumption goods. Hence, the government can simply base the non-linear tax on the total income net of new savings, which by the budget constraint equals consumption expenditures.

When the dual labor market is optimal, the government will step in to provide partial insurance to the fixed-term workers. This public insurance scheme will involve savings distortions implied by the inverse Euler equation.

Lemma 6. *Suppose that in the optimum consumption is bounded below: $\forall_{h\in\mathcal{H}} c(h) \geq \underline{c} > 0$. Then the inverse Euler equation holds*

$$\forall_{h\in\mathcal{H}\setminus\mathcal{H}_T} \frac{1}{u'(c(h))} = \sum_{s\in\mathcal{H}_{|h|+1}(h)} \frac{\mu(s \mid h)}{u'(c(s))}.$$

Golosov et al. [2003] derive this condition in the NDPF setting under a general productivity process and show that it imply a positive tax rate on savings. Kocherlakota [2005] proposes a decentralization that involves linear capital tax with the rate that depends on current and previous labor income level. Golosov and Tsyvinski [2006] interpret this savings distortion as an asset testing. Hence, the fixed-term workers will face a history dependent public insurance system that involves asset tests.

7 Empirical evidence

The endogenous commitment framework of Section 6 yields testable implications about the income risk of labor contracts with different firing cost. In this section I use the administrative data of employment spells in Italy to show that indeed the fixed-term contracts coincide with a higher residual variance of income.

7.1 Labor contracts in Italy

Italy in its modern history experienced a proliferation of distinct labor contracts.[10] I focus on only two types of contract: permanent (*il contratto a tempo indeterminato*) and fixed-term (*il contratto a tempo determinato*). Prior to the reforms in 2014 the permanent contract used to feature exceptionally high firing cost. In the permanent contract, an employer can legally dismiss her worker for two reasons, which essentially are difficult situation of the firm or inadequate fulfillment of tasks by the worker. Any fired worker could sue the company for an unfair dismissal. If the judge decides that dismissal was unfair, the worker had a right to be rehired by the original firm and compensated for the income lost during the legal process.[11] Ichino et al. [2003] provides evidence that judges decision are not impartial: judges were less likely to find a dismissal justified when the unemployment was high. Flabbi and Ichino [2001] suggests that high firing cost leads to very low turnover rate in large Italian service companies.

[10] Tealdi [2011], who provides an overview of labor reforms in Italy, state that in 2006 there were 46 different labor contracts.

[11] As Ichino et al. [2003] put it, 'Virtually, firing costs are higher in Italy than anywhere else, because this is the only country in which, if firing is not sustained by a just cause falling under the above two headings, the firm is always forced to take back the employee on payroll and to pay the full wage that he/she has lost during the litigation period plus welfare contributions; in addition, the firm has to pay a fine to the social security system for the delayed payment of welfare contributions up to 200 percent of the original amount due. '. Reforms in 2015

The fixed-term contracts do not allow for worker's dismissal justified by a difficult situation of the company. However, as the contract expires, the firm may decide not to extend the contract and hence terminate the employment relationship at no cost.[12] I conclude that the permanent and the fixed-term contracts in Italy are a good empirical counterparts of the permanent and fixed-term contracts as described in the framework of Section 6.

7.2 Empirical model

I measure the lack of insurance residually, as a variation in income which cannot be explained by fixed personal characteristics, age, tenure, labor market experience, firm type, sector, location or time effects. Consider the following model, following Guiso et al. [2013]:

$$\log\left(y_{ijt}\right) = \rho + W_{it}'\alpha + F_{jt}'\beta + M_{ijt}'\gamma + D_t'\delta + \epsilon_{ijt}, \tag{10}$$

where

- W_{it} includes workers time invariant and time varying characteristics,
- F_{jt} includes firm characteristics time invariant and time varying characteristics,
- M_{ijt} includes match characteristics, like tenure and type of contract,
- D_t are yearly fixed effects,
- ϵ_{ijt} is the error term.

We want to estimate the variance of ϵ_{ijt}. Let's compute the difference of (10)

$$\log\left(\frac{y_{ijt}}{y_{ijt-1}}\right) = \Delta W_{it}'\alpha + \Delta F_{jt}'\beta + \Delta M_{ijt}'\gamma + \Delta D_t'\delta + \Delta\epsilon_{ijt} \tag{11}$$

Take a vector of variables $X \in \{W, F, M\}$ and denote its vector of parameters by ξ. Divide X into three components:

$$X_{ijt} = [X_{ij}^1, X_{ijt}^2, X_{ijt}^3],$$

where X_{ij}^1 involves variables which are fixed in time, X_{ijt}^2 variables that depend linearly on year, such as age, labor market experience or tenure, and X_{ijt}^3 are variables that depend on time nonlinearly. Let's separate the vector of parameters ξ in the same way into ξ_1, ξ_2 and ξ_3. Then we can write

$$\Delta X_{ijt}'\xi = \sum \xi_2 + \Delta X_{it}^{3\prime}\xi_3.$$

[12] In the period I consider the firm could extend the fixed-term contract once. The second extension lead to the automatic conversion of the contract into a permanent one. Labor reforms in 2014 allowed to up to 5 extensions that together with the original contract last no longer than 3 years.

Hence (11) becomes

$$\log\left(\frac{y_{ijt}}{y_{ijt-1}}\right)=\sum\alpha_2+\sum\beta_2+\Delta W_{it}^{3\prime}\alpha_3+\Delta F_{jt}^{3\prime}\beta_3+\Delta M_{ijt}^{3\prime}\gamma_3+\Delta D_t^{\prime}\delta+\varepsilon_{ijt},$$

where $\varepsilon_{ijt}=\Delta\epsilon_{ijt}$. This way we avoid the need to estimate the fixed effects of workers, firms and match, which greatly reduced the number of parameters. Furthermore, this specification is robust to possible correlation between individual fixed effects and tenure or labor market experience.

I am interested in how the variance of the error term in (10) depends on a particular match characteristic: a type of employment contract. I assume that the distribution of error ϵ_{ijt} is independent of time and I denote the variance of error with the permanent contract by σ_P^2 and the variance of error with the fixed-term contract by σ_{FT}^2. Let's call $\frac{\sigma_{FT}^2}{\sigma_P^2}$ a *risk ratio*. The risk ratio greater than 1 means that the fixed-term contracts imply more income risk, or equivalently less income insurance, than the permanent contracts. The risk ratio is equal

$$\frac{\sigma_{FT}^2}{\sigma_P^2}=\frac{1-\rho_P}{1-\rho_{FT}}\frac{Var\left(\varepsilon^{FT}\right)}{Var\left(\varepsilon^{P}\right)}$$

where ρ_x is the autocorrelation of errors when contract is $x\in\{P,FT\}$. If errors for two contract types have the same autocorrelation, then the risk ratio is simply given by the ratio of variances of errors from the difference equation (11). If the autocorrelation of the permanent contract is greater, than the risk ratio is smaller than the observed $\frac{Var\left(\varepsilon^{FT}\right)}{Var\left(\varepsilon^{P}\right)}$. I estimate the correlation of errors with the following identity

$$\rho_x=\frac{1+\frac{Cov\left(\varepsilon_{t-1}^x,\varepsilon_t^x\right)}{Var\left(\varepsilon^x\right)}}{2+\frac{Cov\left(\varepsilon_{t-1}^x,\varepsilon_t^x\right)}{Var\left(\varepsilon^x\right)}}.$$

7.3 Data

I use *Work Histories Italian Panel* (WHIP), a sample of administrative records of Italian employment histories.[13] The time-span in which permanent and fixed-term contracts can be observed separately is 1997-2004. The data is at the annual frequency. I consider only a full time jobs. I annualize the real income from a given job by dividing it by an average number of working days.

I extract all two-period employment spells of a given individual at the given firm with a contract of a given type. As an illustration of this procedure, consider the following example of a work history.

[13] Work Histories Italian Panel is a database of work histories developed thanks to the agreement between INPS and University of Turin. More information on http://www.laboratoriorevelli.it/whip.

Table 2: Regression of $\hat{\varepsilon}_t^2$ (main estimates)

variable	**coefficient**	**t**	**95% confidence interval**
constant	0.0347***	10.557	$(0.028, 0.041)$
fixed-term contract	0.009***	13.058	$(0.008, 0.01)$
$\log(y_{ijt})$	-0.0019***	-5.591	$(-0.003, -0.001)$

*** - statistically significant at the 1% level.

Table 1: An example of an employment history

year	**company**	**contract**
1998	A	fixed-term
1999	A	fixed-term
2000	B	fixed-term
2001	B	permanent
2002	B	permanent
2003	B	permanent

A worker with such an employment history was working on a fixed term contract for company A for two years. Then the worker moved to a company B for one year of fixed-term employment followed by the permanent employment. From this employment history I extract three two-period employment spells: $1998 : 1999$ at the company A with a fixed term contract and $2001 : 2002, 2002 : 2003$ at the company B with the permanent contract. I do not use the spell $1999 : 2000$, as it involved a change of an employer. I also do not utilize the spell $2000 : 2001$, as it involved a change of contract.

For each 2-period employment spell, I compute the logarithm of ratio of annualized income. I remove outliers separately for two types of contract by considering only the spells with $\log\left(\frac{y_{ijt}}{y_{ijt-1}}\right)$ within three standard deviations from the sample mean. I use the following explanatory variables: worker characteristics (gender, geographical region), firm characteristics (firm's age, sector), match characteristics (tenure, type of job) as well as annual dummies.

7.4 Results

I estimate the equation (11) for each type of contract separately with OLS.[14] Then I take squared residuals from both regressions, pool them into a one vector and regress them on a set of explanatory variables that includes a 'fixed-term contract' dummy variable. This procedure is essentially the White test for heteroskedasticity of the error term (see White [1980]). A significant positive estimate of the parameter of the 'fixed-term contract' dummy means that the fixed-term contracts are associated with higher variance of errors from the difference equation (11). The main results of this regression are reported in Table 2, the full results are reported in Appendix **??**.

The fixed-term dummy is positive and highly significant, which means that the variance of errors of the auxiliary differenced regression are higher for the fixed-term contracts: $Var\left(\varepsilon^{FT}\right) > Var\left(\varepsilon^{P}\right)$.

[14]There are 179,831 two-period spells with the permanent contract and 3,486 with the fixed-term contract.

Table 3: Regression of $\hat{\varepsilon}_{t-1}\hat{\varepsilon}_t$ (main estimates)

variable	**coefficient**	**t**	**95% confidence interval**
constant	0.0031	1.383	$(-0.001, 0.008)$
fixed-term contract	-0.0012	-1.568	$(-0.003, 0.001)$
$\log(y_{ijt})$	-0.0006*	-2.742	$(-0.003, -0.001)$

* - statistically significant at the 10% level.

Since variance of errors vary with other characteristics as well (such as log income, as reported in Table 2), in order compute the lower bound on the risk ratio, I consider a male worker from north-west of Italy in 1998, who starts a job in services at the median income ($\approx 20,000$ euros). I find that

$$\frac{Var\left(\varepsilon^{FT}\right)}{Var\left(\varepsilon^{P}\right)} = 1.78.$$

I use similar method to examine the impact of the type of contract on the autocorrelation of errors. I regress the product of the lagged and current residuals on a set of explanatory variables and a 'fixed-term contract' dummy. In this case the impact of the fixed-term contract is statistically insignificant at 10% level (see Table 3). Moreover, using the point estimate for a median male worker as before, I arrive at the correlation ratio

$$\frac{1-\rho^{P}}{1-\rho^{FT}} = 0.9997.$$

Hence, the risk ratio is very well approximated by the ratio of variances of error from the differenced equation.

$$\frac{\sigma_{FT}^2}{\sigma_P^2} \approx \frac{Var\left(\varepsilon^{FT}\right)}{Var\left(\varepsilon^{P}\right)} = 1.78.$$

The income risk faced by the median worker with fixed-term contract is 78% higher than the income risk faced by the similar worker with the permanent contract. It is an economically significant value. A worker with the permanent contract earning a median income can expect that with 95% probability his next year income will be between $17,509$ and $23,777$ euros. The same worker with a fixed-term contract will have a wider confidence interval of $16,522$ to $24,927$ euros.

The analysis above may suffer from a selection problem. That would be the case if firms offering more risky jobs use fixed-term contracts, while more stable firms hire on the permanent basis. A proper causal analysis of relation between a type of contract and the residual volatility of income is an interesting topic for future research.

8 Quantitative exercise

In this section I calibrate the simple life-cycle model to the Italy and derive the normative prescriptions regarding the optimal income taxation and labor regulation.

Table 4: Estimated income process

Age ↓ \ contract type →	Temporary workers ($\underline{\theta}$)			Permanent workers ($\overline{\theta}$)		
Young	$y(\underline{\theta})$:	16,745		$y(\overline{\theta})$:	20,147	
	$\mu(\underline{\theta})$:	0.53		$\mu(\overline{\theta})$:	0.47	
		Unlucky	**Lucky**		**Unlucky**	**Lucky**
Old	$y(\underline{\theta}, \theta')$:	18,839	25,506	$y(\overline{\theta}, \theta')$:	20,349	26,504
	$\mu((\underline{\theta}, \theta') \mid \underline{\theta})$:	0.51	0.49	$\mu(\overline{\theta}, \theta' \mid \overline{\theta})$:	0.4	0.6

8.1 Calibration

I consider only wage workers with a full-time job. I divide my sample into two age groups: young (below the median age) and old (above or equal to the median age). With the data at hand I cannot observe the persistence of income on such a long time period - at best I observe 8 income years for each individual. Rather than assuming the earning process that is independent across time, I use the data on total employment spell with a given employer. Within each age group, I divide workers between permanently employed (with permanent employment contract and sufficiently long total employment spell at the current employer) and temporarily employed (with shorter total employment spell). I assume that income of permanently employed old is informative about the future income of permanently employed young[15]. Another rationale for this division is that, according to the theory, the data on labor income is more informative of productivity for workers that are not engaged in long-term relationship with their employers. As it turns out, for both types of workers at each history the income is strictly increasing in age (see table 4). Under assumption of no borrowing in the data, the income process is informative about the productivity for all age/contract type groups.[16]

I take the mean labor income of young within each contract group and assign probability of each contract group by relative frequency in the data. For the old, within each contract group I describe the earnings distribution with the Gaussian mixture model. The Gaussian mixtures can approximate well complex distributions (Marin et al. [2005]) and were successfully used to capture higher moments of the US earnings distribution (Guvenen et al. [2015]). I estimate the mixture by maximizing likelihood with the Expectation-Maximization algorithm of Dempster et al. [1977]. Then, in order to keep the model simple, I take the estimated means of each component of the mixture as a distinct earning realization that occurs with the probability equal to the weight of this component in the mixture. In practice, for both groups of old workers (permanently and temporarily employed) the mixture of two normal distributions fits the data well. Table 4 presents the estimation results. Income is reported in euros per year at the 2004 prices.

I use the contemporaneous log utility and disutility of labor that implies the compensated elasticity of labor supply of 1 :

$$u(c) - v(n) = \log(c) - \Gamma \frac{n^2}{2}.$$

[15] In fact, in the dataset there are some permanent workers that cross the age threshold between age groups.

[16] Since I consider only two periods, the upward time trend dominates the stochastic variation. In the future work I plan to estimate the model for more age groups, where the issue of disentangling current output and insurance is likely to emerge.

I am left with 7 parameters to determine: the productivity at each history and the labor disutility parameter Γ. I pin down productivities using the first-order condition of labor supply

$$\theta = y(\theta)\sqrt{\Gamma\frac{1-T(y(\theta))/y(\theta)}{1-T'(y(\theta))}}, \tag{12}$$

where $T(y)$ is the actual tax Italian tax schedule.[17] As the income realizations at each history fall into the same tax bracket, the square root term is roughly the same for at all histories. Since the Italian tax schedule is progressive, the marginal tax rate is greater than the average tax rate and the square root term is slightly greater than 1. It means that estimated productivities are roughly proportional to income, with the constant of proportionality slightly greater than one. Without the loss of generality I set Γ to 1 - according to (12) varying Γ would simply rescale all the productivities. I set the discount factor β to 0.5, roughy corresponding to the period of 17 years.

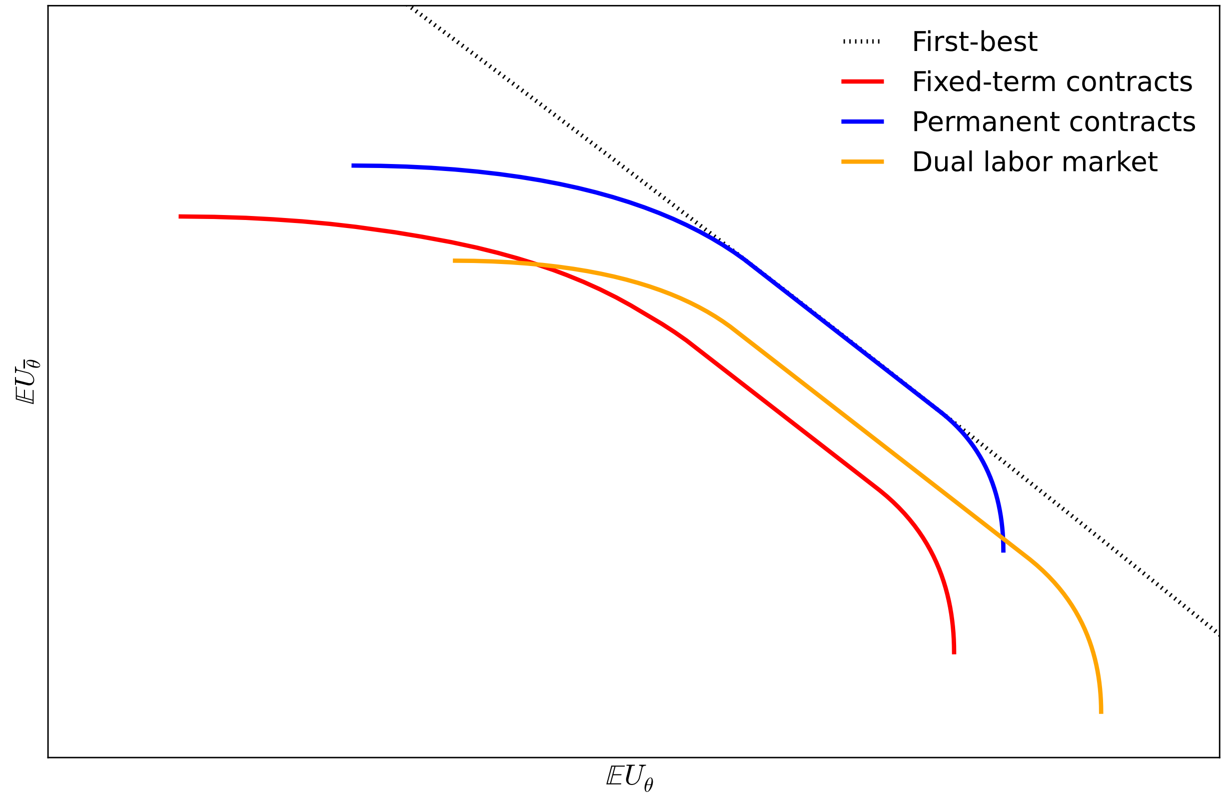
8.2 Pareto Frontiers

In Figure 1 I draw Pareto frontiers of three different regimes. 'Fixed-term contracts' regime corresponds to the NDPF economy, in which all workers receive fixed-term contracts and firms do not engage in insurance provision. 'Permanent contracts' regime is characterized by both initial types receiving permanent employment. Finally, the 'Dual labor market' frontier describes the economy in which the initially more productive type is employed permanently, while the other is employed on a fixed-term basis.[18] I plot as well the Pareto frontier of the first-best allocation as an indicator of what is feasible, even if we abstract from the incentive issues. The first-best is characterized by the full consumption insurance and efficient allocation of labor at each history. In each of these allocations I require the government to raise the same net tax revenue as in the actual Italian tax schedule.

[17] Italy undertook a series of tax reforms in the considered period. I use the tax schedule from year 2000, which captures the average shape of the tax function in these years.

[18] I do not plot the Pareto frontier of the fourth configuration, in which the initially low productivity type receives the permanent contract and the initially high type receives the fixed-term contract, as it is already implicitly incorporated in the 'Permanent contracts' frontier. Under permanent contracts regime with two types the government can still distort labor supply of the initially high type upwards just as if tis type had the fixed term contract - compare Proposition 2 and its proof.

Figure 1: Pareto frontiers with different contract assignments



Whenever the worker is employed on the fixed-term contract, the only source of insurance against the productivity risk is the tax system. Information constraints prevent the government from implementing simultaneously full consumption insurance and efficient allocation of labor. As a result, for any regime where some initial type is employed on the fixed-term basis, the Pareto frontier is bounded away from the first-best. On the other hand, the permanent contracts allow for coexistence of full insurance and efficient labor supply when the redistribution between initial types is limited. Therefore, the permanent contracts frontier coincides with the first-best, when the social preferences are not strongly redistributive.

Further, note that each constrained efficient allocation with only fixed-term contracts is Pareto dominated by at least one of the other regimes. It's a consequence of Proposition 3, which states that it is always optimal to have some workers permanently employed. Hence, the question is whether it is ever optimal to have some workers that are not permanently employed. As we can see in Figure 1, the case with both workers having permanent contracts is optimal unless welfare preferences are strongly skewed towards the initially less productive type. It is consistent with the intuition behind Proposition 4.

Table 5 presents the optimal allocations under different assumptions about the social preferences. Clearly, the libertarian government that cares only about the insurance and do not want to impose a redistributive taxation prefers the permanent contracts.[19] It turns out that in the calibrated

[19] Equivalently, we can specify the libertarian objective function as $\mathbb{E}U$ subject to no redistribution constraint:

Table 5: Optimal allocations for different social welfare functions

Social preferences:	utilitarian	libertarian	Rawlsian	anti-Rawlsian
Objective	$\mathbb{E}U$	$\sum_{\theta\in\Theta_1}\lambda^{lf}(\theta)\,\mu(\theta)\,\mathbb{E}U_\theta$	$\min_{\theta\in\Theta_1}\mathbb{E}U_\theta$	$\max_{\theta\in\Theta_1}\mathbb{E}U_\theta$
Optimal regime	permanent contracts	permanent contracts	dual labor market	permanent contracts
Consumption insurance	full	full	partial	full
Allocation of labor	efficient	efficient	distorted	distorted
Welfare (cons. equiv.)				
laissez-faire = 100%	104.3%	104.1%	110.4%	107.2%
NDPF = 100%	101.5%	101.5%	102.8%	101.2%

economy a utilitarian government that maximizes the average workers' utility would ban fixed-term contracts and introduce no labor distortions as well.[20] The remaining two social welfare functions represent extremes, in which the government cares only about the least or the most well-off type of worker. Both such governments rely strongly on redistribution, hence the labor supply will be distorted.[21] The Rawlsian government prefers to give fixed term contracts to the initially less productive type, as it allows to increase taxation on the top worker. On the other hand, the permanent contracts do not restraint the redistribution of the anti-Rawsian government (see discussion under Proposition 2), hence there is no reason to resort to insurance-inferior fixed-term contracts.

The last two rows of Table 5 compare the constrained efficient allocations with endogenous commitment with the laissez-faire and NDPF allocations in terms of welfare. The welfare metric is the consumption equivalent measure: by which factor shall we increase consumption of workers at each history in the benchmark allocation (laissez-faire or NDPF) to obtain the same welfare as in the constrained efficient allocation. When we consider the less redistributive governments (utilitarian and libertarian), the NDPF captures two-thirds of the gains from constrained efficient allocation. t means that the simpler tax systems that encourages firms' insurance improve upon the complicated tax systems prescribed by NDPF literature by more than 50% in terms of welfare. The relative welfare gain from using the permanent contracts in comparison to the fixed-term contracts is smaller for governments focused on redistribution (Rawlsian and anti-Rawlsian).

9 Conclusions

Firms are crucial intermediaries of information and tax payments between workers and the government. This intermediation should be acknowledged by the optimal tax theory, as it can resolve some of its shortcomings. Farhi and Werning [2013] and Golosov et al. [2015] argue that the complicated,

$\forall_{\theta\in\Theta_1}\sum_{h\in\mathcal{H}(\theta)}R^{1-|h|}T(h)=0$. In other words, the net present value of tax burden of each initial type is 0. All the remaining taxation plays an insurance role and workers would voluntarily decide to participate in such public insurance scheme.

[20] Although the utilitarian optimum lays on the first-best Pareto frontier, it does not coincide with the first-best utilitarian optimum. The government constrained only by the budget constraint would equalize consumption between two types, while keeping the labor supply undistorted. Such allocation is not incentive compatible even with permanent contracts.

[21] Rawlsian case: $\underline{\theta}$ distorted downwards, $\overline{\theta}$ efficient, as well as distortions in the second period for initial type $\underline{\theta}$. Anti-Rawlsian case: $\underline{\theta}$ efficient, $\overline{\theta}$ distorted upwards.

history dependent tax schedules can guide us towards simpler, but still close to optimal taxation. Instead, I show that incorporating firms into the framework leads to the fully optimal tax system that is remarkably simple. Under mild assumptions, most of the individuals face a 'static', history independent redistributive tax. Moreover, if the government values redistribution to the poor, low income workers have access to the means-tested public insurance program. Such an arrangement resembles the actual tax and transfer systems of many countries.

Furthermore, I uncover a novel trade-off in public finance. By empowering the private sector to insure workers, the government has to give up part of its ability to redistribute income. I interpret this finding through the lenses of dual labor market, in which fixed-term workers face volatile income, while permanent workers are insured by their employers. My model suggests that, once we take redistribution into account, the dual labor market may be constrained efficient. It is worth to examine how does the trade-off between redistribution and insurance influence the optimal regulation of other insurance markets.

References

Laurence Ales and Pricila Maziero. Non-exclusive dynamic contracts, competition, and the limits of insurance. 2009.

Pierre-Andre Chiappori, Bernard Salanie, and Julie Valentin. Early starters versus late beginners. *Journal of Political Economy*, 107(4):731–760, 1999.

Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38, 1977.

Peter A Diamond. Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review*, pages 83–95, 1998.

Emmanuel Farhi and Iván Werning. Insurance and taxation over the life cycle. *The Review of Economic Studies*, 80(2):596–635, 2013.

Luca Flabbi and Andrea Ichino. Productivity, seniority and wages: new evidence from personnel data. *Labour Economics*, 8(3):359–387, 2001.

Mikhail Golosov and Aleh Tsyvinski. Designing optimal disability insurance: A case for asset testing. *Journal of Political Economy*, 114(2):257–279, 2006.

Mikhail Golosov and Aleh Tsyvinski. Optimal taxation with endogenous insurance markets. *The Quarterly Journal of Economics*, 122(2):487–534, 2007.

Mikhail Golosov, Narayana Kocherlakota, and Aleh Tsyvinski. Optimal indirect and capital taxation. *Review of Economic studies*, pages 569–587, 2003.

Mikhail Golosov, Aleh Tsyvinski, and Ivan Werning. New dynamic public finance: a user's guide. In *NBER Macroeconomics Annual 2006, Volume 21*, pages 317–388. MIT Press, 2007.

Mikhail Golosov, Maxim Troshkin, and Aleh Tsyvinski. Redistribution and social insurance. 2015.

Luigi Guiso, Luigi Pistaferri, and Fabiano Schivardi. Credit within the firm. *The Review of Economic Studies*, 80(1):211–247, 2013.

Fatih Guvenen, Fatih Karahan, Serdar Ozkan, and Jae Song. What do data on millions of us workers reveal about life-cycle earnings risk? Technical report, National Bureau of Economic Research, 2015.

Benjamin R Handel, Igal Hendel, and Michael D Whinston. Equilibria in health exchanges: Adverse selection vs. reclassification risk. Technical report, National Bureau of Economic Research, 2013.

Milton Harris and Bengt Holmstrom. A theory of wage dynamics. *The Review of Economic Studies*, 49(3):315–333, 1982.

Igal Hendel and Alessandro Lizzeri. The role of commitment in dynamic contracts: Evidence from life insurance. *Quarterly Journal of Economics*, 118(1), 2003.

Mark Huggett, Gustavo Ventura, and Amir Yaron. Sources of lifetime inequality. *The American Economic Review*, 101(7):2923, 2011.

Andrea Ichino, Michele Polo, and Enrico Rettore. Are judges biased by labor market conditions? *European Economic Review*, 47(5):913–944, 2003.

Narayana R Kocherlakota. Zero expected wealth taxes: A mirrlees approach to dynamic optimal taxation. *Econometrica*, 73(5):1587–1621, 2005.

Narayana R Kocherlakota. *The new dynamic public finance*. Princeton University Press, 2010.

Wojciech Kopczuk and Joel Slemrod. Putting firms into optimal tax theory. *The American economic review*, pages 130–134, 2006.

Anna Kosior, Michał Rubaszek, and Kamil Wierus. On the importance of the dual labour market for a country within a monetary union. 2015.

Jean-Michel Marin, Kerrie Mengersen, and Christian P Robert. Bayesian modelling and inference on mixtures of distributions. *Handbook of statistics*, 25(16):459–507, 2005.

James A Mirrlees. An exploration in the theory of optimum income taxation. *The review of economic studies*, pages 175–208, 1971.

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1):205–29, 2001.

Cristina Tealdi. Typical and atypical employment contracts: the case of italy. 2011.

Jonathan P Thomas and Tim Worrall. Limited commitment models of the labour market. *Scottish Journal of Political Economy*, 54(5):750–773, 2007.

Halbert White. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica: Journal of the Econometric Society*, pages 817–838, 1980.

A Proofs

A.1 Proofs from Section 3

Proof of Lemma 1. Suppose that there is an equilibrium contract (r, n) which at some history $h \in \mathcal{H}$ yields non-zero profits: $\mathbb{E}\pi_h \neq 0$. If profits are positive, it means that a competitor could profitably steal the worker by offering contract (r', n'), where $r' = r$, $n'(s) = n(s) - \frac{\mathbb{E}\pi_h}{2\mu(s|h)R^{|h|-|s|}\theta(s)}$. This contract yields $\frac{1}{2}\mathbb{E}\pi_h$ profits and would be preferred by the worker, as it involve less labor supply. Suppose instead that profits are negative. Then the firm would simply terminate the contract. Hence, all equilibrium contracts under no commitment satisfy $\forall_{h\in\mathcal{H}}\mathbb{E}\pi_h = 0$.

Suppose there is an equilibrium contract (r, n) which does not belong to $\arg\max_{r,n}\mathbb{E}U_{\theta_1}(y, a, T, r, n)$, s.t. $\forall_{h\in\mathcal{H}}\mathbb{E}\pi_h = 0$. It means there is another contract $(\bar{r}, \bar{n})$ which yields strictly greater expected utility to the worker subject to the zero profit condition at each history. Define $r' = \bar{r}$ and $\forall_h n'(h) = \bar{n}(h) + \epsilon$, where $\epsilon > 0$. For epsilon sufficiently small (r', n') contract will yield expected utility greater than (r, n). Moreover, this contract yields positive profits. Hence, (r, n) cannot be the equilibrium contract and we know that $\mathcal{C}^{NC} \subseteq \arg\max_{r,n}\mathbb{E}U_{\theta_1}(y, a, T, r, n)$, s.t. $\forall_{h\in\mathcal{H}}\mathbb{E}\pi_h = 0$.

Suppose that there is a contract $(r, n) \in \arg\max_{r,n}\mathbb{E}U_{\theta_1}(y, a, T, r, n)$, s.t. $\forall_{h\in\mathcal{H}}\mathbb{E}\pi_h = 0$ that is not an equilibrium contract. It means that there is another contract (r', n') that yields positive profits and the expected utility greater than (r, n). In turn, this implies that there is yet another contract which yields zero profits at each history and the expected utility greater than (r', n'). It contradicts the fact that (r, n) belongs to the $\arg\max_{r,n}\mathbb{E}U_{\theta_1}(y, a, T, r, n)$, s.t. $\forall_{h\in\mathcal{H}}\mathbb{E}\pi_h = 0$. Hence, we know that $\{\arg\max_{r,n}\mathbb{E}U(y, a, T, r, n)$, s.t. $\forall_{h\in\mathcal{H}}\mathbb{E}\pi_h = 0\} \subseteq \mathcal{C}^{NC}$.

I will show the last claim by induction. Zero profit conditions at the histories of length $\mathcal{T}$ imply $\forall_{h\in\mathcal{H}^{\mathcal{T}}} y(h) = \theta(h)n(h)$. Consider history of length t and suppose that for all histories of length greater than t labor income equals output. Then $\forall_{h\in\mathcal{H}^t}\mathbb{E}\pi_h = \theta(h)n(h) - y(h)$, which is equal to zero by the zero profit condition.□

A.2 Proofs from Section 4

Proof of Lemma 2 The proof follows from the proof of Lemma 1. We just need to note that since workers and firms commit not the terminate the contract, the competition for workers takes place only in the initial period. Hence, the zero profit condition needs to hold only at the initial histories. □

Proof of Theorem 1 Take some incentive-compatible allocation (y, a, T) which does not involve full consumption insurance. For each initial type $\theta_1 \in \times_1$ find a full-insurance consumption level $\bar{c}(\theta_1)$ with equality $\sum_{t=1}^{\mathcal{T}}\beta^{t-1}u(\bar{c}(\theta_1)) = \sum_{h\in\mathcal{H}(\theta_1)}\beta^{|h|}\mu(h)u(c(h))$. Construct a new allocation $(\bar{y}, \bar{a}, \bar{T})$ which implies full consumption insurance at the level $\bar{c}(\theta_1)$ and the same expected life-time income, conditional on initial type report. For instance, it can be done by setting $\bar{y} = y$, $\bar{a} = a$ and $\forall_{h\in\mathcal{H}}\bar{T}(h) = T(h) + c(h) - \bar{c}(h_1)$. The utility of truthful reporting each initial type is unchanged and since the allocation involves more insurance, the government saves some resources. Now I'll show that at full insurance allocation all the other incentive constraints are satisfied.

First, I'll show that full insurance allocation is incentive-compatible if and only if the lifetime labor income along any history depends only on the initial type:

$$\forall_{\theta_1 \in \Theta_1} \forall_{h,s \in \mathcal{H}^{\mathcal{T}}(\theta_1)} \sum_{t=1}^{\mathcal{T}} R^{1-t} y\left(h^t\right) = \sum_{t=1}^{\mathcal{T}} R^{1-t} y\left(s^t\right).$$

Suppose this equality does not hold - lifetime income from history s is strictly greater than the lifetime income from history h, where both h and s start from the same initial type: $s_1 = h_1$. Because of full insurance, consumption allocation is identical in the two allocations. Since one of them involve more income, by the zero profit condition of the firm it also involves higher levels of labor supply. It means that workers are strictly better always reporting h than s, hence the incentive-compatibility is not satisfied. Conversely, if a worker prefers reporting h to s, when both these allocations involve full insurance, it must be that at the history s the worker supplies more labor. Supplying more labor means that the life-time income at the history h is greater.

Second, note that it is always possible to equalize income along all histories. Simply set $y(h)$ to the average of histories of this length: $\sum_{s \in \mathcal{H}_{|h|}(h_1)} \mu(s \mid h)\, y(s)$. Since income becomes deterministic, full consumption insurance can be obtained with asset trades when sufficient borrowing is allowed, or with age-dependent taxes. □

Proof of Lemma 3. I'm going to show that under Assumption 1 the following two properties hold: $V_\theta(C,Y)$ is decreasing in θ and $\frac{\partial}{\partial Y} V_\theta(C,Y)$ is increasing in θ. These two conditions imply that $V_\theta(C,Y)$ satisfies the single crossing condition. In the proof I use the notation as if productivities were continuously distributed. The similar proof holds when the types are discretely distributed, however the notation becomes much more involved.

First I will show that $V_\theta(C,Y)$ is decreasing in θ. Take two initial types $\theta_1, \theta_2 \in \Theta_1$ such that $\theta_2 > \theta_1$. Denote the allocation of labor corresponding to the initial type θ_1 and lifetime income Y by $n_1 : \mathcal{H}(\theta_1) \to \mathbb{R}_+$. We can define an inverse function $M^{-1} : \mathcal{H} \times [0,1] \to \mathbb{R}_+$ that goes from the last period history and a real value between 0 and 1 to the particular value of the new productivity draw, conditional on history h.[22] We can express the allocation of labor n_1 not as a function of the current history, but as a function of last period history h and the value of the cdf: $\tilde{n}_1(h,x) = n_1\left(h, M^{-1}(h,x)\right) \forall_{h \in \mathcal{H}, x \in [0,1]}$. Now suppose we apply the labor allocation $\tilde{n}_1$ for the initial type θ_2: $\tilde{n}_2\left(\theta_2, h_2, ..., h_{|h|}, x\right) = \tilde{n}_1\left(\theta_1, h_2, ..., h_{|h|}, x\right)$. Then the lifetime labor disutility type θ_2 gets from $\tilde{n}_1$ is identical to the lifetime disutility of type θ_1. However, as

$$\int_{\inf \Theta_{|h|}}^{\sup \Theta_{|h|}} \sum_{\theta \in \Theta_{|h|}} \mu(\theta \mid h)\, v\left(\tilde{n}_2(h, M(h,\theta))\right) = \int_{M^{-1}(h,0)}^{M^{-1}(h,1)} \mu(\theta \mid h)\, v\left(\tilde{n}_2(h, M(h,\theta))\right) d\theta = \int_0^1 v\left(\tilde{n}_2(h,x)\right) dx.$$

Note that the expected utility from following $\tilde{n}_2$ for one period after the history h is identical to

[22] I am interested only on the histories with positive density. On the set of these histories the cdf is increasing, so we can define the inverse function.

the expected utility of θ_1 following the labor allocation $\tilde{n}_1$. On the other hand, the expected output of the initial type θ_2 given previous history h is weakly greater for type θ_2

$$\int_{\inf \Theta_{|h|}}^{\sup \Theta_{|h|}} \mu(\theta \mid h)\, \theta \tilde{n}_2(h, M(h,\theta))\, d\theta = \int_0^1 M^{-1}(\theta_2, s, x)\, \tilde{n}_2(\theta_2, s, x)\, dx \geq \int_0^1 M^{-1}(\theta_1, s, x)\, \tilde{n}_1(\theta_1, s, x)\, dx,$$

as $M(\theta_1, s, \theta) \geq M(\theta_2, s, \theta) \implies M^{-1}(\theta_1, s, x) \leq M^{-1}(\theta_2, s, x)$. Note that the expected lifetime output is strictly greater for θ_2, as at least in the initial period this type produces more than θ_1. Hence, the lifetime disutility of labor is identical, but type θ_2 produces more, which means that such an allocation of labor is strictly suboptimal. Optimally, the and θ_2 utility in equilibrium will be higher. Hence, $V_\theta(C, Y)$ is increasing in θ.

Now I will show that $\frac{\partial V_\theta(C,Y)}{\partial Y}$ is increasing with θ. First note that $\frac{\partial V_\theta(C,Y)}{\partial Y} = -\frac{v'(n(h))}{\theta(h)}$ for any $h \in \mathcal{H}(\theta)$. Suppose that $\frac{v'(n(h))}{\theta(h)}$ does not decrease with θ. It means that $\forall_{s,\theta} n(\theta_1, s, \theta) \leq n(\theta_2, s, \theta)$ for $\theta_2 > \theta_1$, where $s \in \times_{i=2}^{|h|-1} \Theta_i$.Then the expected output is

$$\int_{\Theta_1} \mu(\theta \mid \theta_1, s)\, \theta n(\theta_1, s)\, d\theta \leq \int_{\Theta_1} \mu(\theta \mid \theta_2, s)\, \theta n(\theta_2, s)\, d\theta,$$

since the labor supply is increasing with θ and one probability distribution weakly first order stochastically dominates the other. This inequality is strict at least in the initial period, hence the lifetime output of type θ_2 is strictly greater than that of θ_1. It violates the initial zero profit condition of the firm, which cannot happen in equilibrium. □

Proof of Theorem 2 It was shown in the main text that the optimal dynamic taxation problem under full commitment can be stated as the static Mirrlees [1971] taxation problem. The solution to the Mirrlees model, when the first order approach is valid, was expressed in terms of elasticities by Saez [2001]. The first-order approach is valid if the single-crossing condition holds and the resulting income schedule is non-decreasing. Hence, what remains to be shown are the relevant elasticities, which I derive below.

First, let's define $\phi(\theta_1, Y)$ as $\frac{v'(n(h))}{\theta(h)}$ where $h \in \mathcal{H}(\theta_1)$ and $Y(\theta_1) = Y$. Let's define the marginal tax rate $T'(\theta_1)$ as $1 - \frac{\phi(\theta_1, Y(\theta_1))}{u'\left(\left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)^{-1} C(\theta_1)\right)}$. I will derive the elasticities by varying the marginal tax rate.

The compensated elasticity is given by

$$\bar{\zeta}^c = \frac{\partial Y(\theta_1)}{\partial 1 - T'(\theta_1)} |_{dC(\theta_1)=0} \frac{1 - T'(\theta_1)}{Y(\theta_1)} = \frac{\phi(\theta_1, Y(\theta_1))}{\phi_Y(\theta_1, Y(\theta_1))\, Y(\theta_1)}.$$

With an implicit function theorem we can get the derivative of ϕ with respect to income. First, with (8) express $n(h)$ as $g(\theta(h)\, \phi(h_1, Y(h_1)))$, where g is an inverse function of v'. Plug this expression into (9) to get

$$H = \sum_{h \in \mathcal{H}(\theta_1)} R^{1-|h|} \mu(h \mid \theta_1)\, \theta(h)\, g(\theta(h)\, \phi(\theta_1, Y(\theta_1))) - Y(\theta_1) = 0.$$

By the implicit function theorem we have

$$\phi_Y(\theta_1, Y(\theta_1)) = -\frac{\partial H}{\partial Y(\theta_1)}\left(\frac{\partial H}{\partial \phi(\theta_1, Y(\theta_1))}\right)^{-1} = \left(\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)(\theta(h))^2 g'(\theta(h)\phi(\theta_1, Y(\theta_1)))\right)^{-1} =$$

$$= \left(\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)\frac{(\theta(h))^2}{v''(n(h))}\right)^{-1}.$$

Hence, we have

$$\bar{\zeta}^c = Y(\theta_1)^{-1}\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)\frac{\theta(h)\, n(h)\, v'(n(h))}{n(h)\, v''(n(h))} = \sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)\frac{\theta(h)\, n(h)}{Y(\theta_1)}\zeta^c(h),$$

where $\zeta^c(h)$ is the compensated elasticity of labor supply at history h. The lifetime compensated elasticity is the average compensated elasticity across all histories starting in θ_1, weighted by the realized output.

The uncompensated elasticity is given by

$$\bar{\zeta}^u = \frac{\partial Y(\theta_1)}{\partial 1 - T'(\theta_1)}\frac{1 - T'(\theta_1)}{Y(\theta_1)} = \bar{\zeta}^u + \frac{\left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)^{-1} u''\left(\left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)^{-1} C(\theta_1)\right)}{\phi_Y(\theta_1, Y(\theta_1))}(1 - T'(\theta_1)).$$

Denote the wealth effect by $\bar{\xi} = \bar{\zeta}^c - \bar{\zeta}^u$. Then

$$\bar{\xi} = \left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)^{-1}\left(\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)\frac{(\theta(h))^2(1 - T'(\theta_1))\, u''}{v''(n(h))}\right)$$

$$= \left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)^{-1}\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)\,\xi(h),$$

where $\xi(h)$ is the wealth effect at the history h. The lifetime wealth effect is the average wealth effect across all histories.

Given the elasticities , we can express the optimal marginal tax rates with the formula of Saez (2001)

$$\frac{T'(\theta)}{1 - T'(\theta)} = \frac{1 - Mu(\theta)}{\theta\mu(\theta)}\frac{1 + \bar{\zeta}^u}{\bar{\zeta}^c}\mathbb{E}\left\{(1 - \omega(\theta'))\frac{u'(c(\theta))}{u'(c(\theta'))}\middle|\,\theta' \geq \theta\right\}.$$

□

A.3 Proofs from Section 5

Proof of Lemma 4. The proof follows from the proof of Lemma 1. Since firms can commit not the terminate the contract, we do not have to worry about firms' incentives to fire workers after the initial period. As a result, the zero profit condition needs to hold as equality only at the initial histories. □

Proof of Proposition 1. First, I will show that when the borrowing limit b is sufficiently high, absent the government intervention firms provide workers with full insurance. Then I will prove that if the borrowing limit is not sufficiently high, the government can substitute borrowing with taxes.

Suppose that the government does not impose any taxes: $\forall_{h \in \mathcal{H}} T(h) = 0$ and that the borrowing until the last period is unconstrained: $b = +\infty$. Take a structure of labor income and labor supply corresponding to the allocation without commitment: y^{NC}, n^{NC}. Modify it in the following way: for each initial type θ and for any length of the history $t \geq 2$ set $\forall_{h \in \mathcal{H}_t(\theta)} y(h) \equiv \max_{h \in \mathcal{H}_t(\theta)} y^{NC}(h)$. It means that starting from the second period, labor income of the initial type θ is deterministic and follows the maximal income possible in this period under the no commitment allocation. In order to compensate firm for expected losses in periods $t \geq 2$, deduct the expected losses from the initial labor income: $y(\theta) \equiv y(\theta) - \sum_{h \in \mathcal{H}(\theta)} R^{1-|h|} \mu(h \mid \theta) \left(y(h) - y^{NC}(h) \right)$.[23] Set $n \equiv n^{NC}$. Now the firm makes zero profits in the initial period and non-positive expected profits in each period that follows (strictly negative, if productivity remains stochastic after the given history). Hence, all the limited commitment constraints are satisfied. Note that as the labor income is deterministic and borrowing is unconstrained (but has to be repaid), the worker obtains full consumption insurance by trading the risk-free asset. Finally, in order to obtain the equilibrium contract the firm will have to modify the labor supply policy in order to adjust it to the constant consumption. The borrowing required to achieve full consumption insurance is equal to finite and bounded above by $\bar{b} = \max_{h \in \mathcal{H}(\theta)} Y(\theta) - y(\theta)$, where $Y(\theta)$ stands for the expected lifetime labor income $\sum_{h \in \mathcal{H}(\theta)} R^{1-|h|} y(\theta)$.

As the government can obtain full consumption insurance, it is clear that the government can implement any full commitment allocation. Now I will show that the government can implement any full commitment allocation even when the borrowing by itself is not sufficient. Suppose that the workers cannot borrow. Denote by $(y_{\bar{b}}, a_{\bar{b}}, T_{\bar{b}})$ the optimal choice of government when the borrowing limit is $\bar{b}$. Set $y = y_{\bar{b}}$ and a_0 to zero. Set $\forall_{h \in \mathcal{H}} T(h) = T_{\bar{b}} + a_{\bar{b}} - R a_{\bar{b}} \left(h^{-1} \right)$. Then the allocation of consumption and labor supply under no borrowing the government chooses (y, a, T) is identical to the optimum under borrowing limit $\bar{b}$. □

Proof of Proposition 2 . Take any constrained efficient allocation under no commitment $\left(y^{NC}, a^{NC}, T^{NC}, n^{NC} \right)$ in which the incentive constraint with respect to reporting strategy $r_{\overline{\theta}, \underline{\theta}}$ is not binding. First I will show that the same allocation can be implemented under one-sided commitment. Then I will show that the government can improve upon it by a careful provision of insurance.

In this simple case the lack of commitment of workers, conditional on reporting strategy r, implies the following upper bound on labor supply

In the constrained efficient allocation under no commitment the labor supply policy of the initial $\overline{\theta}$ is either undistorted or distorted upwards. Set $y(h) = y^{NC}(h)$, $a(h) = a^{NC}(h)$ and $T(h) =$

[23] If the stochastic volatility of productivity is high, the initial labor income may become negative.

$T^{NC}(h)$ for all $h \in \mathcal{H}(\overline{\theta})$. Under limited commitment the firm in principle could try to modify the labor policy in order to provide the worker with higher utility, as it does under full commitment. Conditional on truthful reporting, the equilibrium contract cannot increase the labor supply in the last period above the level $\frac{y^{NC}(h)}{\theta(h)}$ for any $h \in \mathcal{H}_2$, since otherwise the worker would have incentives to quit:

$$\forall_{h \in \mathcal{H}_2} \mathbb{E}\pi_h \leq 0 \Leftrightarrow \frac{y(r(h))}{\theta(h)} \geq n(h).$$

Hence, the only modification in the labor supply policy the firm can do is to decrease the second period labor supply and, in order not to suffer losses, increase it in the first period. However, as $n^{NC}(\overline{\theta})$ is either undistorted or distorted upwards, while $n^{NC}(h), h \in \mathcal{H}_2(\overline{\theta})$ is either undistorted or distorted downwards, such perturbation would always decrease the welfare of worker. Hence, the government in able to implement exactly the same allocation of $\overline{\theta}$ type under one-sided commitment as under no commitment. Note that the same reasoning applies to the initial $\underline{\theta}$ type, but in this case we know for sure that the initial labor supply policy $n^{FT}(\underline{\theta})$ is undistorted as IC_r does not bind.

Now let's see how the government can improve upon the no commitment allocation. Suppose that giving full commitment allocation (as described in Proposition 1) for both types with the same net present value of taxes is incentive compatible. Since both agents receive full insurance, they are strictly better off.

Suppose that implementing full commitment allocation for both types with the same net present value of taxes makes type $\theta_1 \in \mathcal{H}_1$ want to mimic the other type $\theta_2 \in \mathcal{H}_1 \backslash \theta_1$. Then rather than giving full insurance to θ_2, give this type the no commitment allocation. Then type θ_1 is strictly better of not mimicking θ_2 and the government can slightly increase redistribution from θ_1 to θ_2. This way both types are better off than in the no commitment allocation. $\square$

Proofs from Section 6

Proof of Proposition 3. Suppose that all initial workers are employed at the fixed term contracts. The taxation problem is then equivalent to no commitment case. There is a type θ such that no other type is tempted to report θ (see Lemma 7 below). Give the permanent contract to type θ, keeping the present value of taxes fixed. The utility of this worker increases and the incentive-compatibility constraints of this type are now slack. If the incentive-compatibility constraints of some other type is violated, make a transfer from θ to the other type until incentive-constraints is satisfied as equality. $\square$

Lemma 7. *There exists a type $\bar{\theta} \in \Theta_1$ such that all the incentive-compatibility constraints (7) of all other initial types with respect to reporting initially $\bar{\theta}$ are slack.*

Proof of Lemma 7. First I will show that the directed graph defined by the binding incentive constraints (7) of the optimal allocation does not contain cycles. Then I will prove that the lack of cycles implies that $\bar{\theta}$ exists.

Suppose that there is a vector of initial types $x = (\theta_1, ..., \theta_j, \theta_1) \in \Theta_1^{j+1}$ that constitutes a cycle, i.e. for $i \leq j$ the incentive-compatibility constraint (7) is binding, where $r_i(x_i) = x_{i+1}$. Denote the expected lifetime tax revenue from type θ by $\mathbb{E}T(\theta) \equiv \sum_{h\in\mathcal{H}(\theta)} R^{1-|h|}\mu(h \mid \theta) T(h)$. There always exist an index $i \leq j$ such that the expected tax revenue is not decreasing, going from type x_i to $x_{i+1} : \mathbb{E}T(x_i) \leq \mathbb{E}T(x_{i+1})$. Given that, the government can modify the outcome of the direct mechanism by setting $\forall_{h\in\mathcal{H}(x_i)} (y, a, T)(h) = (y, a, T)(r_i(h))$. The expected utility of type x_i is unchanged, as the incentive constraint w.r.t. r_i was binding. Moreover, the tax revenue from type i does not decrease. Hence, the welfare does not decrease as well. It means that this constraint was not binding in the first place, since the government can get rid of this constraint at no cost. I conclude that in the optimal allocation there are no cycles of binding incentive constraints.

Now, suppose that for every initial type θ there exists another type θ' such that the incentive constraint w.r.t. r', where $r'(\theta') = \theta$, is binding. Denote the set of such types by $P(\theta) = \{\theta' \in \Theta_1 :$ incentive constraint w.r.t. r' is binding, where $r_i(\theta') = \theta\}$. We are assuming that this set is non-empty for any θ. Consider the following algorithm.

1. Take any element θ of the set Θ_1. Take an initially empty set X.
   (a) If $P(\theta) \cap X \neq \varnothing$, break - there is a cycle.
   (b) Else, assign $X \equiv X \cup P(\theta)$.
   (c) For each $\theta' \in P(\theta)$, assign $\theta \equiv \theta'$ and go to the step (b).

Note that in each iteration the set X either increases, or the loop is stopped. The loop is stopped if the same initial type was encountered twice, which indicates the cycle. Since the set of all initial types Θ_1 is finite, at some point the same type has to be met twice and the loop is stopped. Hence, under the initial assumption made, there must be a cycle. Conversely, if there is no cycle, then there exist some type $\bar{\theta}$ such that no other type is tempted to mimic him. Since in the optimal allocation there are no cycles, such type always exists. $\square$

Proof of Proposition 4. First let's suppose that $\underline{\theta} = 0$, so that $n(\underline{\theta}) = 0$ in each case. The Rawlsian government maximizes the transfers given to the workers of type $\underline{\theta}$. Since workers are risk-neutral, there will be no insurance-driven taxation in the second period. First, suppose that $\underline{\theta}$ has the permanent contract. The binding incentive compatibility constraint is

$$Y\left(\bar{\theta}\right) - T_P\left(\bar{\theta}\right) - \sum_{h\in\mathcal{H}\left(\bar{\theta}\right)} \beta^{|h|-1}\mu\left(h \mid \bar{\theta}\right) v\left(n\left(h\right)\right) = Y\left(\underline{\theta}\right) + \frac{\mu\left(\bar{\theta}\right)}{\mu\left(\underline{\theta}\right)} T_P\left(\bar{\theta}\right) - \sum_{h\in\mathcal{H}\left(\underline{\theta}\right)} \beta^{|h|-1}\mu\left(h \mid \bar{\theta}\right) v\left(\tilde{n}\left(h\right)\right),$$

where $\tilde{n}$ is the unique labor policy of deviating type $\bar{\theta}$ which minimizes the disutility cost of producing $Y(\underline{\theta})$

$$\tilde{n} \equiv \arg \min_{n:\mathcal{H}(\underline{\theta})\to\mathbb{R}_+} \sum_{h\in\mathcal{H}(\underline{\theta})} \beta^{|h|-1}\mu\left(h \mid \bar{\theta}\right) v\left(n\left(h\right)\right) \text{ s.t. } \sum_{h\in\mathcal{H}(\underline{\theta})} R^{1-|h|}\mu\left(h \mid \bar{\theta}\right) \theta\left(h\right) n\left(h\right) \geq Y\left(\underline{\theta}\right).$$

On the other hand, if type $\underline{\theta}$ receives the fixed-term contract and produces the same output $Y(\underline{\theta})$, the incentive constraint is

$$Y\left(\overline{\theta}\right)-T_{FT}\left(\overline{\theta}\right)-\sum_{h\in\mathcal{H}\left(\overline{\theta}\right)}\beta^{|h|-1}\mu\left(h\mid\overline{\theta}\right)v\left(n\left(h\right)\right)=Y\left(\underline{\theta}\right)+\frac{\mu\left(\overline{\theta}\right)}{\mu\left(\underline{\theta}\right)}T_{FT}\left(\overline{\theta}\right)-\sum_{h\in\mathcal{H}\left(\overline{\theta}\right)}\beta^{|h|-1}\mu\left(h\mid\overline{\theta}\right)v\left(\bar{n}\left(h\right)\right),$$

where he labor policy of deviating type $\overline{\theta}$ is just $\bar{n}$ is just $\bar{n}\left(\overline{\theta}\right)=\frac{\underline{\theta}}{\overline{\theta}}n\left(\underline{\theta}\right)=0$ and $\bar{n}\left(\left(\overline{\theta},\theta\right)\right)=n\left(\left(\underline{\theta},\theta\right)\right)\forall_{\theta\in\Theta_2}$.

The difference in redistribution between the two cases is

$$T_{FT}\left(\overline{\theta}\right)-T_{P}\left(\overline{\theta}\right)=\mu\left(\underline{\theta}\right)\left(\sum_{h\in\mathcal{H}\left(\overline{\theta}\right)}\beta^{|h|-1}\mu\left(h\mid\overline{\theta}\right)v\left(\bar{n}\left(h\right)\right)-\sum_{h\in\mathcal{H}\left(\underline{\theta}\right)}\beta^{|h|-1}\mu\left(h\mid\overline{\theta}\right)v\left(\tilde{n}\left(h\right)\right)\right).$$

Since $\tilde{n}$ is a unique policy that minimizes disutility cost of producing $Y(\underline{\theta})$, this difference is strictly positive: $T_{FT}\left(\overline{\theta}\right)>T_{P}\left(\overline{\theta}\right)$. Note that in general the type $\underline{\theta}$ will produce a different amount under the fixed-term contract.[24] However, it does not matter for the proof. I showed that even with sub-optimal choice of labor income, this type $\underline{\theta}$ can gain on the transition from the permanent to fixed-term contract.

What is different when $\underline{\theta}>0$? The reasoning above holds when we compare allocations in which the initial labor supply of type $\underline{\theta}$ is undistorted. However, the government may wish to distort downwards to relax the incentive-compatibility constraint. It is unclear how efficient are distortions in the two considered regimes: it may be the case that distortions harm type $\underline{\theta}$ less under permanent contracts. In such a case, if $\underline{\theta}$ is sufficiently large, the government will have a large scope for applying distortions (undistorted $n(\underline{\theta})$ will be small), the permanent contracts for $\underline{\theta}$ may be desirable. However, if the scope for distortions is sufficiently small ($\underline{\theta}$ small), the gains from distortions will be small as well and will not eradicate the advantage of the fixed-term contract. □

Proof of Proposition 5. Under the permanent contract, consumption of type $\underline{\theta}$ in each period is constant and equal $-T(\underline{\theta})$. If we give this type the fixed-term contract without changing the allocation of labor supply, the utility of $\underline{\theta}$ will not change. There is no consumption risk, since the worker has the same (zero) labor income at each history. However, now the government can increase the utility of $\underline{\theta}$ by lifting distortions. Since productivity in the second period is independent of the productivity in the first one, the only reason to distort second period labor of $\underline{\theta}$ is insurance in the second period. Specifically, it means that the highest productivity type should be undistorted.

To see it, note that under fixed-term contract the utility $\overline{\theta}$ gets from misreporting initial type is equal to utility of the initial $\underline{\theta}$. Originally we are in the situation

$$\mathbb{E}U_{\overline{\theta}}=\mathbb{E}U_{\underline{\theta}}.$$

[24]Since productivity in the second period is independent of the first period draw, there is no gain from distorting labor supply of type $\underline{\theta}$ in the second period under the fixed-term contract. On the other hand, under permanent contracts the labor supply will be distorted downwards, as type $\overline{\theta}$ can use the superior first period productivity to produce the second period income.

Suppose that we can increase utility of $\underline{\theta}$ in a budget-neutral way, simply by reducing distortions. Denote the gain from reducing distortions by $\Delta > 0$. Crucially, note that both $\underline{\theta}$ and the mimicking type $\overline{\theta}$ gain exactly the same utility by reducing distortions. It leads to

$$\mathbb{E}U_{\overline{\theta}} < \mathbb{E}U_{\underline{\theta}} + \Delta,$$

so the incentive constraint is violated. In order to keep it satisfied, let's change the net present value of taxes of each type in a way that does not affect the budget balance

$$\mathbb{E}U_{\overline{\theta}} + \Delta^T_{\overline{\theta}} = \mathbb{E}U_{\underline{\theta}} + \Delta^D + \Delta^T_{\underline{\theta}}.$$

Now note that the total impact on $\underline{\theta}$ is $\Delta^D + \Delta^T_{\underline{\theta}}$, which is just equal to $\Delta^T_{\overline{\theta}}$. Moreover, since we change the taxes in a budget neutral way, $\Delta^T_{\overline{\theta}} > 0$. Hence, the type $\underline{\theta}$ strictly gains on lifting distortions, even if we take into account a decrease in redistribution. $\square$

Proof of Proposition 6. Suppose that the tax depends only on current consumption expenditures: $T_x\left((x_i)_{i=1}^t\right) \equiv \bar{T}(x_t)$. Take worker at the initial history θ. He may deviate from the optimum of having $\forall_{s\in\mathcal{H}(\theta)} x(s) = Y(\theta)\left(\sum_{t=1}^{\mathcal{T}} R^{1-t}\right)$ by either choosing another constant level of expenditures or by introducing some volatility to expenditures. Note that the first type of deviation is taken care of by the incentive-compatibility constraints from the original problem. We just have to extend $\bar{T}_x$ such that it is defined for all levels of expenditures and that the tax on values of x that are not claimed by any initial type in the direct mechanism is sufficiently high to deter deviation.

Introduction of volatility is more problematic. Suppose that a worker decides to have a sequence of expenditures $(x_t)_{t=1}^{\mathcal{T}}$. Suppose that $\bar{T}_x$ is convex. Then the resulting lifetime consumption is $\sum_{t=1}^{\mathcal{T}} R^{1-t}c_t = \sum_{t=1}^{\mathcal{T}} R^{1-t}\left(x_t - \bar{T}_x(x_t)\right) \leq \sum_{t=1}^{\mathcal{T}} R^{1-t}\left(\bar{x} - \bar{T}_x(\bar{x})\right)$, where $\bar{x} = \frac{\sum_{t=1}^{\mathcal{T}} R^{1-t}x_t}{\sum_{t=1}^{\mathcal{T}} R^{1-t}}$. Note that the labor supply allocation is the same when expenditures are $(x_i)_{i=1}^t$ and when they are $(\bar{x})_{i=1}^t$, as the total lifetime income is the same. Hence, the utility from this deviation is bounded above by the utility from deviating to have a constant expenditures $\bar{x}$. However, as we have seen in the previous step, deviations to have a constant income are already taken care of.

Suppose instead that $\bar{T}_x$ is not convex. Then there will be some sequence of expenditures that a worker will be able to increase the average consumption at the expense of some volatility. For sufficiently risk-neutral agents such deviations will be profitable. $\square$

Proof of Lemma 6. The proof follows the intuition from Golosov et al. [2003]. Take some history $h \in \mathcal{H}\backslash\mathcal{H}_{\mathcal{T}}$ and consider a small perturbation δ

$$c'(h) = c(h) + \frac{\delta}{u'(c(h))},\ \forall_{s\in\mathcal{H}_{|h|+1}(h)} c'(s) = c(s) - \frac{\delta}{\beta u'(c(s))},$$

and c' elsewhere is set to c. Note that utility from no reporting strategy has changed. It means that all the incentive compatibility constraints hold. In order for such perturbation not to be beneficial, it has to be the case that it yields no free resources

$$-\frac{\delta\mu\left(h\right)}{u'\left(c\left(h\right)\right)}+\sum_{s\in\mathcal{H}(h)_{|h|+1}}\frac{\mu\left(s\mid h\right)\delta}{\beta u'\left(c\left(s\right)\right)}=0,$$

which implies that the inverse Euler equation holds. $\square$