Optimal Taxation with Permanent Employment Contracts

Paweł Doligalski*

November 16, 2015

[the latest version]

Abstract

New Dynamic Public Finance describes the optimal income tax in an environment where private insurance is absent. I extend this framework by introducing permanent employment contracts which facilitate insurance provision within firms. The optimal tax system becomes remarkably simple, as the government outsources most of the insurance provision to employers and focuses mainly on redistribution. When the government wants to redistribute to the poor, a dual labor market could be optimal. Less productive workers are hired on a fixed-term basis and are partially insured by the government, while the more productive ones enjoy the full insurance provided by permanent employment. I provide empirical evidence consistent with the theory and characterize the constrained efficient allocations for Italy.

1 Introduction

Lifetime incomes differ due to initial heterogeneity in earning potential of workers and luck experienced during the working life.[1] The standard welfare criteria call for the elimination of both types of inequality. New Dynamic Public Finance (NDPF) answers this call by designing a tax system that both redistributes income between initially different people and insures them against differential luck realizations.[2] This approach has been criticized for two reasons. First, it neglects private insurance possibilities. Second, the optimal tax system is far more complicated than any tax system observed in reality. In this paper I address these two problems of NDPF by introducing permanent employment contracts.

The individual productivity of each worker evolves as a random process. Insuring a worker essentially means keeping his consumption constant through times of both high and low productivity.

*European University Institute, pawel.doligalski[at]eui.eu. I am grateful for valuable comments of Árpád Ábrahám, Juan Dolado, Piero Gottardi, David Levine, Ramon Marimon, Dominik Sachs and seminar participants in the National Bank of Poland, European University Institute and Warsaw Economic Seminar. I thank the members of EUI Writers' Group for useful language advice. All mistakes are mine.

[1] Huggett, Ventura, and Yaron (2011) estimate that out of the two, initial differences account for more than 60% of the inequality in lifetime earnings.

[2] Golosov, Tsyvinski, and Werning (2007) and Kocherlakota (2010) survey the NDPF literature.

Insurance via income tax is difficult because the government does not observe individual productivity.[3] I assume that firms have better information than government, yet face a different friction: neither they nor workers are able to commit to maintain the employment relationship. Permanent contracts enable employers to credibly commit to not firing workers, thus allowing them to act as insurers.[4] The government optimally outsources most of the insurance to the better informed firms and, depending on the social objectives, can focus on redistribution. As a result, the optimal tax system is much simpler and resembles the taxes used in reality. In the model calibrated to Italy any constrained efficient allocation can be implemented with a tax schedule that depends exclusively on current consumption expenditure. It contrasts with the standard implementation of NDPF which involves time-varying tax on labor income and capital income that depends on the history of past earnings.

There is a cost to promoting insurance within firms. Permanent contracts reduce the random variation of income over the life-cycle, but they also allow firms to structure worker's compensation in a way that minimizes the worker's tax burden.[5] Such a behavior reduces the government's ability to redistribute. A redistributive government sometimes prefers to strip the least productive workers of the private insurance by equipping them with fixed-term contracts, since in this way they either receive higher transfers or face lower labor distortions. Hence, I provide a novel rationale for the coexistence of permanent and fixed-term contracts.

In the model economy risk averse workers face risk due to stochastic idiosyncratic productivity and can trade only a risk-free asset. Risk neutral firms observe workers' productivity and compete for them in the labor market. The labor market is frictional, as both parties are unable to commit to maintain the employment relationship in the future. Workers are free by law to change employers. Firms can, at a specified cost, fire employees. I consider two different types of labor contract: permanent and fixed-term. Fixed-term contracts allow firms to dismiss workers in every period without any cost. Permanent contracts have high firing cost which discourages firms from laying off their workforce.

When all workers have fixed-term contracts, the taxation problem is equivalent to NDPF. If a firm and a worker can terminate their relationship at no cost and start a new one with a clean slate, no private insurance is possible. Worker's income is equal to his output in each period and the labor market collapses to a sequence of spot labor markets. Optimally, the government steps in with taxation that both redistributes and insures. Since the government is constrained by available information, it has to set up a complicated, history dependent income tax system to screen evolving productivities of workers. Golosov, Kocherlakota, and Tsyvinski (2003) show that the optimal insurance provision with private information requires levying a tax on labor income and on savings, although agents are heterogenous only in labor productivity.

When all workers are employed on a permanent basis, firms are not tempted to fire workers, but workers are unable to commit to stay in their firms. I show that this market imperfection can be remedied by backloading labor compensation. By shifting labor income to the future, employers

[3] Financial markets are unlikely to insure workers for the same reason.

[4] Guiso, Pistaferri, and Schivardi (2005) document that Italian firms insure their workers by reducing the variability of their income.

[5] Best (2014) shows that firms adjust wage offers to the details of the workers' tax system.

effectively lock workers in the company. As workers no longer have incentives to quit, firms can offer them full consumption insurance. In such cases the taxation problem can be expressed as a static Mirrlees (1971) model.[6] The tax schedule depends on the average lifetime elasticity of labor supply and only the initial distribution of types. Intuitively, if all people entered 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.

I show that the government should always assign permanent contracts to some workers. Specifically, efficiency requires that workers who pay the highest taxes should be fully insured and permanently employed. The intuition is simple: with permanent contract, paying high taxes becomes more attractive. It could, nevertheless, be suboptimal to equip all workers with permanent contracts. When the government cares most about the initially least productive, these workers could optimally end up with fixed-term contracts and no private insurance. The reason behind this finding is as follows. Under permanent contracts firms can shift workers' income to the future. On the one hand, this allows firms to insure workers; on the other, firms have incentives to structure income payments in a way that minimizes their employees' tax burden. The currently productive workers would benefit from shifting income to the future and claiming transfers due to low current earnings. Since such income shifting is possible only under permanent contract, the government can prevent this by assigning fixed-term contracts at low levels of income. This argument provides a novel perspective on dual labor markets where the two types of contracts coexist, a prevalent labor market arrangement in Europe. There is the extensive literature documenting the negative impact of dual labor markets on the unemployment risk, the human capital accumulation and the volatility of business cycles.[7] I complement this literature by highlighting the additional redistributive channel of fixed-term contracts.

How to implement the optimal allocation with taxes? When all workers have permanent contracts, they should face only the redistributive tax based on consumption expenditures. The usual base for redistributive tax, such as labor income or total income, exhibits time variation due to backloading of compensation. Since the consumption expenditures remain stable through a worker's lifetime, it allows the tax schedule to be time-invariant. When tax payments increase progressively with consumption expenditures, the tax schedule can depend only on current consumption expenditures - no history dependence is required. There is no need to tax the savings of permanent workers. When the dual labor market is optimal, fixed-term workers are covered by an extensive public insurance program. As in NDPF, it involves a tax on savings that can be interpreted is as means tested income support.

This paper focuses on the relation between the type of contract and the volatility of workers' income. I show that this effect is present in the data by analyzing the administrative records of employment histories from Italy. The residual income variance of a median worker is higher by 78% under fixed-term rather than permanent contract. This estimate is conditional on continuous employment at one firm, so it is not affected by income changes due to losing or switching jobs.

[6] However, it is not the case that the optimum with only permanent contracts always involves full consumption insurance - see discussion in Section 4.

[7] See references in the related literature section. For information on dual labor markets in Europe, see Eichhorst (2014).

I am the first to document the impact of fixed-term contracts on income volatility, conditional on staying employed. A proper causal analysis of the link between firing costs and income risk is an interesting topic for future research.

I calibrate a simple life-cycle model to Italy. All constrained efficient allocations involve assigning permanent contracts to all workers. As a result, any allocation at the Pareto frontier can be implemented with a simple consumption expenditure tax. The welfare gains are substantial: when the planner is utilitarian, permanent contracts increase welfare gains from optimal taxation by 50%.[8] Then I investigate under which parameter values the dual labor market would be optimal. If the productivity of the initially least productive type was lower by at least 4%, the Rawlsian planner would assign fixed-term contract to these workers.

Related literature. This paper contributes to the literature on optimal taxation with private insurance markets. Golosov and Tsyvinski (2007) study this question under the assumption that the government and firms face the same friction: asymmetric information. I assume that frictions faced by firms and those faced by the government are different: the government lacks information, while firms and workers lack commitment. Stantcheva (2014) considers an environment in which firms face both limited information and limited commitment, but her model is static and hence concerned only with redistribution. Chetty and Saez (2010) model private insurance in the reduced form. Instead, my paper provides microfoundations of insurance on the labor market, which reveals the crucial role of the firing cost. Attanasio and Rıos-Rull (2000) and Krueger and Perri (2011) study how the public insurance crowds out the private one. Although their private insurance is also constrained by the limited commitment friction, agents' endowments are random and exogenous. In my framework productivity is random, but income is endogenous. Shifting income across time turns out to be the key margin of response to taxes. In a different framework Abraham, Koehne, and Pavoni (2014) show that hidden asset trades, which are closely related to income shifting, reduce the optimal progressivity of the labor income tax, which is in line with my findings.

Another strand of the literature focuses on simple tax implementations. Albanesi and Sleet (2006) show that with iid productivity shocks the constrained efficient allocations in NDPF can be implemented with potentially time-varying tax that depends jointly on current wealth and current labor income. Farhi and Werning (2013) and Weinzierl (2011) argue that age dependent taxation captures most of the welfare gains from the optimal non-linear taxes. Findeisen and Sachs (2015) optimize with respect to the history-independent, non-linear labor income tax and linear capital income tax rate. Conesa, Kitao, and Krueger (2009) is an example of a Ramsey approach, which restricts the tax function to some exogenously chosen class. My paper shows that the inclusion of private insurance leads to the fully optimal tax systems that are as simple as the tax functions assumed in the Ramsey approach.

Dual labor markets and fixed-term contracts are studied extensively. It was shown that temporary contracts are associated with higher unemployment risk (García-Pérez, Marinescu, and Castello

[8] Suppose that utilitarian welfare of laissez-faire allocation is 100 in consumption equivalent terms. NDPF achieves 102.8, while optimal taxation with permanent contracts 104.3 (see Table 5). The permanent contracts regime improves NDPF relative to the laissez-faire by more than 50%.

(2014)) and lower on the job training (Cabrales, Dolado, and Mora (2014)) than permanent contracts. Furthermore, dual labor markets amplify macroeconomic fluctuations, as employers are less likely to hoard labor (Bentolila, Cahuc, Dolado, and Le Barbanchon (2012); Kosior, Rubaszek, and Wierus (2015)). I contribute to this literature by documenting that, conditional on continuous employment at one company, fixed-term workers have significantly more volatile income than permanent employees.

The labor market in my model is frictional, as both parties can terminate the contract at any time. There is a long tradition of modeling labor market without commitment, dating back at least Harris and Holmstrom (1982) and Thomas and Worrall (1988). Thomas and Worrall (2007) provide a recent review of the limited commitment models of labor market. This friction plays a key role also in other insurance markets: life insurance (Hendel and Lizzeri (2003)) and health exchanges (Handel, Hendel, and Whinston (2013)).

Structure of the paper. The next section introduces the environment and sets up the taxation problem. The benchmark case without frictions on the labor market is solved in Section 3. Section 4 characterizes the constrained efficient allocation with limited commitment. Implementation with the tax system is discussed in Section 5. In Section 6 I validate the predictions of the model with Italian data. The model is calibrated to Italy in Section 7. The last section concludes. All proofs are available in the Appendix.

2 Framework

2.1 Firms and workers

There is a continuum of workers that live for $\bar{t} \in \mathbb{N}_+$ 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 a 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. A 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 the 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 $1 \leq t \leq \bar{t}$) 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 $\Theta \equiv \cup_{t=1}^{\bar{t}}\Theta^t$. 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 the 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: $\forall_t \sum_{s\in\Theta^t}\mu(s) = 1$. In practice, I will work mostly on the collections of histories that happen with positive probability, denoted by $\mathcal{H} \equiv \{h \in \Theta : \mu(h) > 0\}$. $\mathcal{H}_t$ is the set of histories that happen with positive probability of length t. By $\mathcal{X}(h)$, where $\mathcal{X}$ is a set of histories and $h \in \mathcal{H}$, I denote 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

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

I assume that each initial type faces productivity risk: $\forall_{\theta\in\Theta_1}\forall_{h\in\mathcal{H}_{\bar{t}}(\theta)}\ \mu(h \mid \theta) < 1$.

2.3 Labor market

I make the following assumptions about the labor market. Workers enter the market after their initial productivity is drawn and firms offer them labor contracts (which will be specified below). I assume no search friction, so all workers see all the offers immediately, which leads to a Bertrand competition between firms for workers. Firms observe each worker's productivity, so details of a labor contract are contingent on a worker's history.

Even after the initial contract has been signed, workers are free to leave their current employer and start a new job elsewhere. I impose no mobility cost. On the other hand, firms can fire their workers in any period subject to the specified firing cost.[9] The firing cost is going to be an important policy instrument. I restrict the firing cost, denoted by f, to belong to the set $\left\{0, \bar{f}\right\}$, where $\bar{f}$ is set sufficiently high such that no firm would ever be tempted to fire the worker. I will use the firing cost to distiguish between the *permanent workers* (those for which $f = \bar{f}$) and the *fixed-term workers* ($f = 0$).

[9]One can think about the firing cost as a severance payment to the fired worker. In the setting I consider such interpretation plays no role, as no firing is going to happen in equilibrium.

2.4 Allocation and payoffs

Definition 1. The allocation (c, y, n) specifies consumption $c : \mathcal{H} \to \mathbb{R}_+$, labor income $y : \mathcal{H} \to \mathbb{R}$ and labor supply $n : \mathcal{H} \to \mathbb{R}_+$ at each history.

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

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

I denote the expected utility from the *ex ante* perspective by $\mathbb{E}U(c, n) \equiv \sum_{\theta \in \times_1} \mu(\theta) \mathbb{E}U_\theta(c, n)$. Profits from hiring a worker at the history h given the allocation (c, y, n) are

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

2.5 Direct mechanism

I assume that the social planner observes consumption c, labor income y and the firing cost f, 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.

Definition 2. A *direct mechanism* $(\mathcal{H}, (c, y, f))$ consists of the message space $\mathcal{H}$ and the outcome functions (c, y, f), each going from $\mathcal{H}$ to a relevant subset of $\mathbb{R}$.

The planner sets up a direct mechanism which in each period collects type reports of workers and assigns them consumption, a labor income and a firing cost. As the government does not observe individual labor supply, it is determined in the equilibrium corresponding to the government's choice of outcomes, which I describe below.

The pure reporting strategy r is a function from the set of possible histories to the message space: $r : \mathcal{H} \to \mathcal{H}$. I impose the consistency condition: $\forall_{s,h \in \mathcal{H}} s \in \mathcal{H}(h) \implies r(s) \in \mathcal{H}(r(h))$. It means that consecutive history reports cannot be at odds with which histories are in fact possible. Let's denote the set of consistent pure reporting strategies by $\mathcal{R}$. The truthful reporting strategy r^* is an identity: $r^*(h) = h$ for all $h \in \mathcal{H}$. I allow for mixed reporting strategies $\sigma \in \Delta_{\mathcal{R}}$, where σ is a probability distribution over the pure reporting strategies.

The expected utility of a worker at the history h, given outcome functions (c, y) and a pure reporting strategy r, is $\mathbb{E}U_h(c \circ r, n)$, where $c \circ r$ is a composite function of reporting strategy and consumption function: $(c \circ r)(h) = c(r(h))$. Similarly, the firm's profits are $\mathbb{E}\pi_h(y \circ r, n)$. The reporting strategy affects the outcomes that are assigned by the mechanism. It does not affect labor supply n, as it is decided in equilibrium by firms that observe worker's type. I allow for mixed reporting strategies. The truthful mixed reporting strategy σ^* is a probability distribution concentrated at the truthful pure reporting strategy: $\sigma^*(r^*) = 1$. The payoffs from the mixed reporting strategy $\sigma \in \Delta_{\mathcal{R}}$ given history h are

$$\sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}U_h(c\circ r,n) \quad \text{and} \quad \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}\pi_h(y\circ r,n).$$

2.6 Equilibrium

Firms compete by offering labor contracts to the workers.

Definition 3. A *labor contract* (n, σ) consists of the labor supply allocation n and the reporting strategy σ.

A typical labor contract specifies labor supply and labor income. 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 in the mechanism, constitute a map from type to a labor income. Given y, firms compete by offering the *effective labor income* $y \circ r$.

Lemma 1. *The set of equilibrium contracts is*

$$\mathcal{E}\left(c,y,f\right) \equiv \arg \max_{\substack{n\,:\,\mathcal{H}\to\mathbb{R}_+\\ \sigma\,\in\,\Delta_{\mathcal{R}}}} \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}U\left(c\circ r,n\right),$$

subject to the zero profit condition

$$\forall_{\theta\in\Theta_1} \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}\pi_\theta\left(y\circ r,n\right) = 0$$

and the limited commitment constraints

$$\forall_{\substack{r\,\in\,\mathcal{R}\\ s.t.\ \sigma(r)>0}} \forall_{h\in\mathcal{H}} - f\left(r\left(h\right)\right) \leq \mathbb{E}\pi_h\left(y\circ r,n\right) \leq 0.$$

Since workers observe all offers, the competition between firms for workers drives profits to zero. The zero profit condition means that firms cannot redistribute. Any transfer of resources between initial types would mean that the firm is making profit on one type and losses on another. It cannot happen in equilibrium, as the profitable type would be captured by the competing firm.

The equilibrium contract has to satisfy the limited commitment constraints: neither worker nor firm can have incentives to terminate the contract at any history. For firms, this means that the expected losses from continuing the employment relationship cannot exceed the firing cost. For workers it means that at no history the firm can make positive profits on them. If that happens, a competing firm could offer them a better deal, while still being profitable.

The mechanism (c, y, f) *implements* allocation (c, y, n) if $(n, \sigma^*) \in \mathcal{E}(c, y, f)$, that is when the labor supply allocation n and a truthful reporting strategy consititute an equilibrium contract.

2.7 The planner's problem

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

$$\max_{\substack{c\,:\,\mathcal{H}\to\mathbb{R}_+\\ y\,:\,\mathcal{H}\to\mathbb{R}\\ f\,:\,\mathcal{H}\to\{0,\bar{f}\}}} \sum_{\theta\in\times_1} \lambda(\theta)\,\mu(\theta)\,\mathbb{E}U(c,n)\,, \tag{3}$$

where λ are the Pareto weights with expected value equal to 1: $\sum_{\theta\in\times_1}\lambda(\theta)\,\mu(\theta)=1$. The optimization is subject to the resource constraint

$$\sum_{h\in\mathcal{H}} R^{1-|h|}\mu(h)\,(y(h)-c(h))\geq 0 \tag{4}$$

and the equilibrium constraint

$$(n,\sigma^*)\in\mathcal{E}(c,y,f)\,. \tag{5}$$

The equilibrium constraint means that there is an equilibrium contract that involves truthful reporting and that the labor supply allocation of workers is determined by this contract. This constraint incorporates the usual incentive compatibility constraints that prevent misreporting. In general $\mathcal{E}(c,y,f)$ is not a singleton and there will be equilibrium contracts that involve misreporting. These other contracts correspond to the binding incentive constraints. I assume that if there exists a truthful equilibrium contract, agents will choose it in equilibrium.[10]

3 Frictionless labor market

In this section I solve the government problem under assumption of private sector operating without frictions: both workers and firms can commit to the labor contract.[11]

Corollary 1. *Under full commitment, the set of equilibrium contracts is*

$$\mathcal{E}^{FC}(c,y)\equiv\arg\max_{\substack{n\,:\,\mathcal{H}\to\mathbb{R}_+\\ \sigma\,\in\,\Delta_{\mathcal{R}}}} \sum_{r\in\mathcal{R}}\sigma(r)\mathbb{E}U(c\circ r,n)\,,\ \ s.t.\ \forall_{\theta\in\Theta_1}\sum_{r\in\mathcal{R}}\sigma(r)\mathbb{E}\pi_\theta(y\circ r,n)=0.$$

Since firms and workers can credibly commit not to terminate the employment relationship, we can drop the limited commitment constraints. This means that the firing cost does not influence the equilibrium. The initial zero profit condition becomes the sole constraint in determination of the equilibrium contract.

In equilibrium the firm chooses the labor supply policy that minimizes the disutility cost of working conditional on satisfying the zero profit condition. The necessary and sufficient conditions for the

[10] It is a usual assumption in the literature. Without it, the planner's problem could have no solution. Note that payoffs of workers and firms are identical for any contract in $\mathcal{E}(c,y,f)$.

[11] Recall that I do not allow firms and workers to contract before the initial productivity draw. I do not consider this a friction, but rather a fundamental element of the economy.

labor supply equalize the marginal cost of an additional unit of output at each history, for each initial type

$$\forall_{h\in\mathcal{H}} \frac{v'(n(h))}{\theta(h)} = \frac{v'(n(h_1))}{\theta(h_1)}. \tag{6}$$

Note that the output and income of a worker do not have to coincide at each history. They have to agree only in expectations over the lifetime of the worker, which is captured by the zero profit condition. Under full commitment firms can smooth labor supply of workers in order to produce the expected lifetime income at the minimal disutility cost. This labor smoothing will be a key element in subsequent analysis.

Let's define the indirect utility function that depends only on the initial type θ, lifetime consumption C and lifetime labor income Y:

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

where the allocation of labor n satisfies labor smoothing (6) and the zero profit condition:

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

Theorem 1. *Under full commitment on the labor market, the planner's problem can be expressed as*

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

subject to the resource constraint

$$\sum_{\theta\in\Theta_1} \mu(\theta)\,(Y(\theta) - C(\theta)) \geq 0$$

and the initial incentive-compatibility constraints

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

By Theorem 1 the planner's problem under full commitment is just the static taxation problem studied by Mirrlees (1971). The planner optimally outsources the entire insurance provision to firms and as a result workers enjoy full consumption insurance. Since firms smooth the labor of workers, only the initial type and the net present value of lifetime income matter for the determination of a worker's allocation of labor. Hence, we can express a worker's utility with an indirect utility function $V_\theta(C,Y)$. As workers need to report only their initial type, only the initial incentive-compatibility constraint is relevant and the entire taxation problem becomes static. The labor market with full commitment provides a microfoundation for interpreting the framework of Mirrlees (1971) as a model of lifetime income taxation.

Let's define the lifetime tax as $T(\theta) \equiv Y(\theta) - C(\theta)$. We can express the optimal marginal tax T'

with the well-known Saez (2001) formula.

Assumption 1. *For any length of the history t and any non-negative number k, if $\theta_2 > \theta_1$ then* $\sum_{s \in \mathcal{H}_t(\theta_1)} \mathbb{I}_{k \geq \theta(s)} \mu(s \mid \theta_1) \geq \sum_{s \in \mathcal{H}_t(\theta_2)} \mathbb{I}_{k \geq \theta(s)} \mu(s \mid \theta_2)$.

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

Proposition 1. *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*

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

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.

Assumption 1 ensures that the indirect utility function $V_\theta(C, Y)$ has the single crossing property.[12] In order to apply the existing optimal tax formulas, we need to slightly modify the environment - Assumption 2 makes the initial distribution of types continuous. Given these assumptions and the additional monotonicity condition, we can express the optimal marginal tax rates with the formula derived by Saez (2001). The optimal 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 a 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 compensated elasticity of labor supply over all histories weighted by output.

4 Frictional labor market

In this section I characterize the optimal allocation when the labor market is frictional: workers can leave firms and firms can fire workers. In this environment the type of labor contract matters, since the high firing cost prevents the firm from terminating the employment relationship in the future. The following two subsections describe the optimal allocation when the planner uses only fixed-term or only permanent contracts. Finally I describe the optimal choice of contract type.

In this section only pure reporting strategies are considered. It is without the loss of generality by Lemma A.2 (in the Appendix).

[12]The single crossing property intuitively means that the higher initial type is more eager to work than the lower initial type. If this property holds, any incentive-compatible lifetime income schedule is non-decreasing in type.

4.1 Only fixed-term contracts

Suppose that the planner assigns fixed-term contracts to workers at each history.

Lemma 2. *Under fixed-term contracts, in any equilibrium contract* (r, n) *at any history* $h \in \mathcal{H}$ *the worker's labor income is equal to the worker's output:* $y(r(h)) = \theta(h) n(h)$.

The zero firing cost under fixed-term contracts means that neither firm nor worker can commit. This lack of commitment implies that neither of the parties can owe any resources to another, as such a loan would never be repaid. As a result, the labor market becomes a sequence of spot labor markets. A worker at each history is paid exactly his current output.

Corollary 2. *Under fixed-term contracts, the planner's problem is a New Dynamic Public Finance taxation problem.*

Lemma 2 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 (5) as

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

This is exactly the incentive-compatibility constraint considered by NDPF.

The solution to the NDPF problem has been studied extensively. As the planner is limited by information, the consumption insurance is only partial. Golosov, Kocherlakota, and Tsyvinski (2003) show that workers' consumption evolves according to the inverse Euler equation, which implies a downward distortion of savings. More recently Farhi and Werning (2013) and Golosov, Troshkin, and Tsyvinski (2015) provide the detailed characterization of the optimal labor wedges.

4.2 Only permanent contracts

Suppose that the planner uses only permanent contracts. The firing cost in this case is assumed to be so high that no firm is ever tempted to fire a worker. Since workers are still free to leave the firm, the labor market operates under one-sided lack of commitment. The equilibrium in this setting was characterized by Harris and Holmstrom (1982) and Krueger and Uhlig (2006).[13] The firm overcomes a worker's commitment problem by backloading labor compensation, i.e. shifting it to the future. As the reward for work comes in the later periods, workers have less incentives to leave the employment relationship early.

Theorem 2. *Take any allocation of consumption and labor supply that can be implemented under the frictionless labor market. The planner can implement it under permanent contracts.*

[13] In Harris and Holmstrom (1982) a firm and a worker learn symmetrically about the worker productivity. They receive noisy signals and the contract is based on the posterior mean of productivity. As the posterior mean is a random variable, this model is equivalent to the framework considered in this paper, where the productivity is observable, but stochastic. Krueger and Uhlig (2006) analyze risk-sharing contracts between risk neutral intermediaries and risk averse agents with risky endowments.

This is one of the main results of this paper. Although the labor market is frictional, as workers cannot credibly promise to stay with their employers, the planner still can provide workers with full consumption insurance.[14] The reasoning is simple due to a direct mechanism approach. The utility of workers depends on their allocation of consumption and not on the allocation of labor income. The limited commitment constraints, on the contrary, depend on labor income but not on consumption. This means that the limited commitment constraints can be relaxed by backloading labor income without affecting the consumption allocation.

We can understand these results in the following way. The firm offers a labor income that is increasing in tenure and varies only with the initial productivity realization. This contract will satisfy the limited commitment constraints, as the labor income is backloaded. The initial compensation can be adjusted such that the firm makes no losses in expectations. Given that the compensation is deterministic, workers can smooth their consumption perfectly by borrowing against future labor income. If the required borrowing is not available due to the borrowing limit, the consumption can be smoothed with age-dependent taxation.

Although the planner can implement full consumption insurance, it will not always be desirable to do so, even when all workers have permanent contracts. In the next subsection I discuss cases in which the planner optimally assigns different types of contracts to different workers, essentially stripping some of them of insurance. Under some circumstances such a dual labor market allocation can be implemented even if all types are nominally assigned permanent contracts.[15] For an example of such a situation, see Lemma A.3 in the Appendix.

4.3 Optimal assignment of contracts

In this subsection I study which type of contracts the planner should assign to workers.

Definition 4. *Top taxpayers* are types that belong to $\arg\max_{\theta\in\mathcal{H}_1}\sum_{h\in\mathcal{H}(\theta)} R^{1-|h|}\mu(h\mid\theta)\left(y(h)-c\left(h\right)\right).$

Theorem 3. *In the optimum, the top taxpayers are assigned permanent contracts and full consumption insurance.*

Assigning permanent contracts allows the planner to provide more insurance, but it also increases the incentives of other workers to misreport. However, there are some types that can be mimicked without loss for the planner: top taxpayers. If any other worker decides to report that he is a top taxpayer, he ends up contributing more resources to the planner's budget. Hence, it is always optimal to assign permanent contracts to top taxpayers. Note that top taxpayers need not be top earners. If the planner cares only about the most productive types, the least productive workers are taxed the most and they should receive permanent contracts.

[14] Harris and Holmstrom (1982) showed that workers can receive full consumption insurance when sufficient borrowing is available (see their footnote 5). My result is more general, as it holds irrespectively of the workers' borrowing limit.

[15] Dual labor market allocation are preferable, because fixed-term contract prevents labor smoothing of the deviating type. However, in some cases the limited commitment constraints of worker are enough to prevent the labor smoothing. That is the case when the deviating worker wants to work less in the first period and more in the second. For details, see Lemma A.3.

Theorem 3 leads us to a strong conclusion: it is never optimal to assign fixed-term contracts to all workers. The planner can always Pareto improve upon the NDPF allocation by introducing permanent employment contracts.

Corollary 3. *If the planner does not want to redistribute between initial types, all workers are optimally assigned permanent contracts and full consumption insurance.*

In the particular case of no redistribution all initial types are top taxpayers. Such a planner cares only about insurance, which can be augmented with permanent contracts.

To study the optimal labor contract of other workers, let's take some allocation (c_0, y_0, n_0) with the contract assignment f_0, where some type $\underline{\theta}$ is assigned permanent contract $f(\underline{\theta}) = \bar{f}$. For simplicity let's assume that the planner places all the welfare weight on $\underline{\theta}$. Suppose that the planner assigns this type fixed-term contract instead and finds the best allocation given the new contract assignment. Denote the new allocation as (c', y', n') and the new contract assignment as f', where $f'(\underline{\theta}) = 0$ and $f'(\theta) = f_0(\theta)$ for any other type. We can decompose the welfare change from the switch of type of contract into three components, capturing the impact on insurance, redistribution and efficiency. [16]

Definition 5. The consumption allocation c_1 solves $\max_{c:\mathcal{H}(\underline{\theta})\to\mathbb{R}_+} \mathbb{E}U_{\underline{\theta}}(c, n_0)$ subject to $\sum_{s\in\mathcal{H}(\underline{\theta})} R^{1-|s|}\mu(s)(c(s) - c_0(s)) = 0$ and $\forall_{r\in\mathcal{R}\cap\mathcal{H}(\underline{\theta})\to\mathcal{H}(\underline{\theta})}\mathbb{E}U_{\underline{\theta}}(c, n_0) \geq \mathbb{E}U_{\underline{\theta}}(c \circ r, n_0)$.

The consumption allocation c_2 solves $\max_{\substack{c\,:\,\mathcal{H}\to\mathbb{R}_+ \\ y\,:\,\mathcal{H}\to\mathbb{R}_+}} \mathbb{E}U_{\underline{\theta}}(c, n_0)$ subject to

$\forall_{s\in\mathcal{H}(\underline{\theta})} y(s) = \theta(s)n(s)$, $\sum_{s\in\mathcal{H}(\underline{\theta})} R^{1-|s|}\mu(s)(c(s) - \theta(s)n(s)) = 0$ and $(r^*, n) \in \mathcal{E}(c, y, f')$.

Define the following three terms:

- $\Delta^{insurance} \equiv \mathbb{E}U_{\underline{\theta}}(c_1, n_0) - \mathbb{E}U_{\underline{\theta}}(c_0, n_0)$,
- $\Delta^{redistribution} \equiv \mathbb{E}U_{\underline{\theta}}(c_2, n_0) - \mathbb{E}U_{\underline{\theta}}(c_1, n_0)$,
- $\Delta^{efficiency} \equiv \mathbb{E}U_{\underline{\theta}}(c', n') - \mathbb{E}U_{\underline{\theta}}(c_2, n_0)$.

The three components sum up to the total welfare impact of the reform:

$$\mathbb{E}U_{\underline{\theta}}(c', n') - \mathbb{E}U_{\underline{\theta}}(c_0, n_0) = \Delta^{insurance} + \Delta^{redistribution} + \Delta^{efficiency}.$$

$\Delta^{insurance}$ captures the welfare loss due increased variability of consumption, keeping the net present value of consumption at the same level as under permanent contracts. $\Delta^{redistribution}$ represents the gain from increased redistribution from other types to $\underline{\theta}$. The transition to fixed-term contract relaxes incentive-compatibility constraints of other types for two reasons: the consumption of $\underline{\theta}$ is more volatile and they cannot use labor smoothing after deviation. The gain from closing the gap in incentive-compatibility constraints is $\Delta^{redistribution}$. Finally, $\Delta^{efficiency}$ corresponds to additional

[16] Doligalski and Rojas (2015) use similar decomposition of welfare change into redistribution and efficiency components in the static Mirrlees (1971) framework with an informal sector.

welfare gains due to simultaneous adjustment of consumption and labor supply along the binding incentive-compatibility constraints. Note that only at this stage the adjustments to labor supply of type $\underline{\theta}$ are allowed.

Lemma 3. $\Delta^{insurance} \leq 0, \Delta^{redistribution} \geq 0, \Delta^{efficiency} \geq 0.$

By Lemma 3 each of the decomposition terms can be signed: the switch from permanent to fixed-term contract leads to a utility loss due to lower insurance and utility gain due to increased redistribution and efficiency. Hence, there are two channels that may lead to the optimal dual labor market: redistribution and efficiency. Consider the following example.

Example 1. There are two initial productivity levels: $\overline{\theta} > \underline{\theta}$. The productivity distribution in periods $t \geq 2$ is independent of the initial type.

The planner wants to maximize the utility of $\underline{\theta}$ and hence will redistribute from $\overline{\theta}$ to $\underline{\theta}$. What limits redistribution is ability of $\overline{\theta}$ to mimic the other type.[17] If the low type has fixed-term contract, the mimicking is straightforward: $\overline{\theta}$ has to produce at each contingency as much as $\underline{\theta}$. If the low type has permanent contract, misreporting will also involve changing the allocation of output. $\overline{\theta}$ is more productive initially and hence will produce more in the first period. Then the firm pays him a part of the first period output as the second period income, allowing the mimicking worker to reduce the second period labor supply. This labor smoothing implies that the high type gains more from misreporting when the other type has permanent contract rather than fixed-term contract.

The following two propositions show two cases under which it is optimal to assign to the initial low type fixed-term contract. Note that by Theorem 3 the high productivity worker will always receive permanent contract.

Proposition 2. *Consider Example 1. Additionally assume that: (i) workers are risk-neutral: $U(c,n) = c - v(n)$, (ii) under permanent contracts the initial type $\underline{\theta}$ has positive lifetime earnings: $\max_{h\in\mathcal{H}(\underline{\theta})} y_0(h) > 0$. The initial type $\underline{\theta}$ can strictly gain by having fixed-term contract.*

The strong assumption of risk neutrality allows us to examine the planner that cares only about the redistribution and not about the insurance ($\Delta^{insurance} = 0$). Since there is no utility loss from volatile consumption, the change in redistribution following the switch of contracts results only from the prevented labor smoothing. We can express the redistribution gain explicitly:

$$\Delta^{redistribution} = \mu\left(\overline{\theta}\right) \sum_{s\in\mathcal{H}(\overline{\theta})} R^{1-|s|}\mu(s \mid \overline{\theta})(v(\tilde{n}'(s)) - v(\tilde{n}_0(s))) \;\; > 0,$$

where $\tilde{n}_0$ is labor allocation of the mimicking type when $\underline{\theta}$ has permanent contract, while $\tilde{n}'$ is the labor allocation when $\underline{\theta}$ has fixed-term contract. $\Delta^{redistribution}$ is strictly positive because of the labor smoothing. Moreover, it is proportional to the fraction of initially productive types. As $\mu\left(\overline{\theta}\right)$ increases, the ratio of taxpayers to transfers receivers increases, allowing the planner to increase the transfer per recipient.

[17] Because of the independence assumption, it is enough to focus only on single deviations from truthful reporting.

Proposition 3. *Consider Example 1. Additionally assume that under permanent contracts the initial type $\underline{\theta}$ has no earnings: $\forall_h \in \mathcal{H}(\underline{\theta}) y_0(h) = 0$. The initial type $\underline{\theta}$ can strictly gain by having fixed-term contract.*

In Proposition 3 it is assumed that the distortions under permanent contracts are so severe that the low type has zero lifetime earnings.[18] Notice that $\Delta^{insurance}$ is zero once again, although this time we do not impose risk neutrality. Since $\underline{\theta}$ does not supply labor supply, there is no need for volatile consumption. Moreover, $\Delta^{redistribution}$ is zero as well, for there is no scope for labor smoothing. The low type can gain only though the efficiency considerations.

Under permanent contracts, the planner discourages misreporting of the high type by reducing labor income of the low type at all histories. It is the case, because the output produced with the superior initial productivity can be paid to him at any history that follows. When the low type has fixed-term contract such income shifting is not possible. Since productivities are independent across time, only distortions of the initial labor income of $\underline{\theta}$ discourage the high type from misreporting. Hence, the planner can lift some of the second period distortions of the low type and achieve $\Delta^{efficiency} > 0$. Specifically, the labor supply at the highest second period productivity will be undistorted.[19] With a fixed-term contract, the low type will have positive earnings and strictly greater utility than with permanent contract.

5 Fiscal implementation

This section is concerned with implementing the optimum with an indirect mechanism - a tax system. The tax can depend on all observables: history of labor income, asset trades and type of contract as well as age.

Definition 6. A *tax system* T is a collection of functions $T = \left\{ T_t \left((y_k, a_k, f_k)_{k=1}^t \right) \right\}_{t=1}^{\bar{t}}$, where

$$T_t : \left(\mathbb{R} \times [-b, \infty) \times \left\{ 0, \bar{f} \right\} \right)^t \to \mathbb{R}.$$

Given the tax system, we can define the set of equilibria.

Lemma 4. *The set of equilibria given the tax system T is*

$$\tilde{\mathcal{E}}(T) \equiv \arg \max_{\substack{c\,:\,\mathcal{H} \to \mathbb{R}_+ \\ y\,:\,\mathcal{H} \to \mathbb{R} \\ n\,:\,\mathcal{H} \to \mathbb{R}_+ \\ a\,:\,\mathcal{H} \to [-b, \infty) \\ f\,:\,\mathcal{H} \to \{0, \bar{f}\}}} \mathbb{E} U(c, n),$$

subject to the zero profit condition

$$\forall_{h \in \mathcal{H}_1} \mathbb{E} \pi_h (y, n) = 0,$$

[18] Optimum under permanent contracts has this feature when $v'(0) > 0$ and $\mu(\underline{\theta})$ is sufficiently low.

[19] It is a consequence of 'no distortion at the top', which holds after each history because productivity draws are independent across time. Note that not all distortions will be lifted, since they serve insurance purpose as well.

the limited commitment constraints

$$\forall_{h\in\mathcal{H}} - f(h) \leq \mathbb{E}\pi_h\,(y,n) \leq 0,$$

the sequence of budget constraints

$$\forall_{h\in\mathcal{H}_1}\; c(h) = y(h) - a(h) - T_1\,(y(h), a(h), f(h))\,,$$

$$\forall_{h\in\mathcal{H}\setminus\mathcal{H}_1}\; c(h) = y(h) + Ra(h^{-1}) - a(h) - T_{|h|}\left(\left(y(h^t), a(h^t), f(h^t)\right)_{t=1}^{|h|}\right),$$

and no borrowing in the terminal period: $\forall_{h\in\mathcal{H}_{\bar{t}}}\; a(h) \geq 0$.

The tax system T *implements* the allocation (c, y, n) if there exist functions a and f such that $(c, y, n, a, f) \in \tilde{\mathcal{E}}(T)$. The main result of this section concerns the fiscal implementation of allocation (c, y, n) in which all workers have permanent contracts and enjoy full consumption insurance.

Definition 7. *Consumption expenditure* at the history h is $x(h) \equiv y(h) + Ra(h^{-1}) - a(h)$. *Consumption expenditure tax* is $\bar{T}_x \equiv \bar{T} \circ \bar{y}^{-1}$, where $\bar{y}(\theta) = \left(\sum_{t=1}^{\bar{t}} R^{1-t}\right)^{-1} \sum_{t=1}^{\bar{t}} R^{1-t} y(h^t)$ for any $h \in \mathcal{H}_{\bar{t}}(\theta)$ is the normalized net present value of labor income of an initial type θ and $\bar{T}(\theta) \equiv \bar{y}(\theta) - c(\theta)$. The *extended consumption expenditure tax* T_x is equal to $\bar{T}_x$ for arguments from the inverse image of $\bar{y}$ and otherwise takes a prohibitively high value.

Theorem 4. *Take any allocation (c, y, n) that is implemented by some direct mechanism and involves full consumption insurance of all workers. Suppose that the borrowing limit is sufficiently high: $b \geq -\min_{h\in\mathcal{H}}(\sum_{t=1}^{|h|} R^{1-|h|}(y(h^t) - \bar{y}(h_1)))$. If $\bar{T}_x$ is convex, then the allocation can be implemented with the tax system*

$$\forall_{t=\{1\ldots\bar{t}\}}\; T_t\left(\left(y_k, a_k, f_k\right)_{k=1}^{t}\right) = T_x\,(x_t)\,.$$

If T_x is not convex, the allocation can be implemented with the tax system

$$\forall_{t=\{1\ldots\bar{t}\}}\; T_t\left(\left(y_k, a_k, f_k\right)_{k=1}^{t}\right) = T_x\,(x_t) + \alpha\,(x_t - x_0)^2\,,$$

where α is high enough such that T_t is convex in x_t.

By Theorem 4 in the simplest case all we need to for fiscal implementation is the redistributive tax based only on current consumption expenditures.[20] Firms offer workers backloaded, but deterministic labor compensation. Workers borrow against it and perfectly smooth consumption. Since consumption expenditures are constant across time for each worker, it is an ideal base for a simple redistributive tax. Note that when the consumption expenditure tax is locally regressive (T_x is not convex), we need to add a corrective term that discourages variation in consumption expenditures. Otherwise some workers could have incentives to vary in their expenditures to reduce the

[20] It is recognized in the literature that a non-linear consumption tax is difficult to implement if the government does not observe each individual transaction. 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. Note that the tax code in US has this feature, as the capital gains are taxed only when they are realized, i.e. when they cease to be savings.

net present value of taxes paid. Conversely, when the consumption expenditure tax is progressive, no history dependence is required.

Corollary 4. *Take any allocation (c, y, n) that is implemented by some direct mechanism and involves full consumption insurance of all workers. Suppose the borrowing is insufficient: $b < -\min_{h \in \mathcal{H}}(\sum_{t=1}^{|h|} R^{1-|h|}(y(h^t) - \bar{y}(h_1)))$. The allocation can be implemented with the tax system as in Theorem 4 combined with the governmental lending to workers.*

Note that in this setting nothing prevents the government from lending to workers. As the debt repayment is contingent only on time, the government can always enforce repayment with taxation. Hence, the government in this setting can always relax the borrowing constraint of the workers enough for Theorem 4 to hold. Cole and Kocherlakota (2001) found the similar relaxation of the borrowing limits to be optimal in the hidden income model with private storage.

Kocherlakota (2005) showed that NDPF can be implemented with the labor income tax that depends on the whole history of labor income and capital income tax which depends on current and previous labor income. Albanesi and Sleet (2006) provide a simpler implementation in the environment with independently and identically distributed productivity shocks. Their tax depends jointly of current income and assets. In these implementations taxes are allowed to vary with time period, or equivalently age of a worker. By Theorem 4 the fiscal implementation of the optimum can be made still simpler with permanent contracts. The tax schedule is time invariant, depends only on current total income net of new savings (with the possible correction term if the tax schedule is regressive) and no additional tax on capital is required. This result holds irrespective of the persistence of productivity shocks. Although assigning permanent contracts to all workers is not always optimal, this the feature of utilitarian optimum in the calibrated model I consider in Section 7.

When the dual labor market is optimal, the fiscal implementation is more complicated, as the tax system both redistributes and insures. Specifically, it has to involve capital taxation of fixed-term workers.

Proposition 4. *Take any allocation (c, y, n) that is implemented by some direct mechanism and involves dual labor market, where consumption of fixed-term workers is bounded away from zero. Fiscal implementation requires taxing assets of fixed-term workers.*

This results is an implication of the inverse Euler equation, which holds in this environment, and the volatile consumption of fixed-term workers. By discouraging savings, capital taxation helps to provide incentives for hard work in the next periods. Following Golosov and Tsyvinski (2006) we can interpret this result as public insurance program with assets testing.

6 Empirical evidence

The model 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 fixed-term contracts coincide with a higher residual variance of income (conditional on continuous employment) and that this difference is economically significant. In a related study Guiso, Pistaferri, and Schivardi (2005) show that Italian firms insure their workers, but they do not differentiate between different types of contracts. Lagakos and Ordonez (2011) document that high-skilled workers receive more insurance within the firm than low-skilled in US. It is consistent with the evidence that high-skilled workers are more frequently offered severance pay (Parsons (2002)).

6.1 Labor contracts in Italy

Italy in its modern history experienced a proliferation of distinct labor contracts.[21] 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 permanent contract used to feature exceptionally high firing cost. An employer could legally dismiss permanent workers for two reasons: 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.[22] Ichino, Polo, and Rettore (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.

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.[23] I conclude that permanent and fixed-term contracts in Italy are close empirical counterparts of permanent and fixed-term contracts as described in the theoretical framework.

6.2 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:

$$\log\left(y_{ijt}\right) = \rho + W_{it}'\alpha + F_{jt}'\beta + M_{ijt}'\gamma + D_t'\delta + \epsilon_{ijt}, \tag{8}$$

where W_{it} includes worker's time invariant and time varying characteristics, F_{jt} includes firm's time invariant and time varying characteristics, M_{ijt} includes match characteristics such as tenure and

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

[22] As Ichino, Polo, and Rettore (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'.

[23] In the period I consider the firm could extend 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.

type of contract, D_t are yearly fixed effects and ϵ_{ijt} is the error term. The parameter of interest is the variance of ϵ_{ijt}, which captures the residual income risk, conditional on being continuously employed. Let's compute the difference of (8)

$$\log\left(\frac{y_{ijt}}{y_{ijt-1}}\right) = \Delta W_{it}'\alpha + \Delta F_{jt}'\beta + \Delta M_{ijt}'\gamma + \Delta D_t'\delta + \Delta\epsilon_{ijt} \tag{9}$$

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 non-linearly. Let's separate the vector of parameters ξ in the same way into ξ_1, ξ_2 and ξ_3. Then we can write

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

and equation (9) becomes

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

where $\varepsilon_{ijt} = \Delta\epsilon_{ijt}$. In this way we avoid the need to estimate the fixed effects of workers, firms and a 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.[24]

How does the variance of the error term in (8) depend on a type of employment contract? Assume that the distribution of error ϵ_{ijt} is independent of time and denote the variance of error with permanent contract by σ_P^2 and the variance of error with 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 fixed-term contracts imply more income risk, or equivalently less income insurance, than 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 differenced equation (9).

[24] See discussion in Guiso, Pistaferri, and Schivardi (2013).

6.3 Data

The data comes from *Work Histories Italian Panel* (WHIP), a sample of administrative records of Italian employment histories.[25] 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 and 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.

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 extract three two-period employment spells are extracted: $1998 : 1999$ at the company A with a fixed term contract and $2001 : 2002, 2002 : 2003$ at the company B with permanent contract. I do not use the spell $1999 : 2000$, as it involved a change of an employer, nor the spell $2000 : 2001$, as it involved a change of contract.

For each 2-period employment spell, the logarithm of ratio of annualized income is computed. 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. The explanatory variables used are: worker characteristics (gender, geographical region), firm characteristics (firm's age, sector), match characteristics (tenure, type of job) as well as annual dummies.

6.4 Results

Equation (9) is estimates with OLS separately for each type of contract.[26] 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 (1980) test for heteroskedasticity of the error term. A significant positive estimate of the parameter of the 'fixed-term contract' dummy means that fixed-term contracts are associated with higher variance of errors from the difference equation (9). The main results of this regression are reported in Table 2, the full results and auxiliary estimates are reported in Appendix B.

[25] 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.

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

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 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.

The fixed-term dummy is positive and highly significant, which means that the variance of errors of the auxiliary differenced regression are higher for fixed-term contracts: $Var\left(\varepsilon^{FT}\right) > Var\left(\varepsilon^{P}\right)$. 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, 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). In this case $Var\left(\varepsilon^{FT}\right)/Var\left(\varepsilon^{P}\right) = 1.78$.

I use similar method to examine the impact of the type of contract on the autocorrelation of errors. The product of the lagged and current residuals is regressed on a set of explanatory variables and a 'fixed-term contract' dummy. In this case the impact of fixed-term contract is statistically insignificant at 10% level (see Table 3). Moreover, using the point estimate for a median male worker as before, we 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 permanent contract. It is an economically significant value. A worker with 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 used fixed-term contracts, while more stable firms hire on a 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.

7 Quantitative exercise

In this section I calibrate the simple life-cycle model using the Italian data (WHIP) and describe the set of constrained efficient allocations.

7.1 Calibration

The sample is divided into two age groups: young (below the median age) and old (above or equal to the median age). Only wage workers with a full-time job are considered. With the data at hand the persistence of income on such a long time period is not observed - at most 8 years of income for each individual are available. 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 (fixed-term workers and workers with shorter total employment spell). I assume that income of permanently employed old is informative about the future income of permanently employed young[27]. 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.[28]

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. The earnings distribution of old workers is described with the Gaussian mixture model. The Gaussian mixtures can approximate well complex distributions (Marin, Mengersen, and Robert (2005)) and were successfully used to capture higher moments of the US earnings distribution (Guvenen, Karahan, Ozkan, and Song (2015)). I estimate the mixture by maximum likelihood Expectation-Maximization algorithm of Dempster, Laird, and Rubin (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 logarithmic utility from consumption and iso-elastic disutility from labor with compensated elasticity of 1:

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

There are 7 parameters left to determine: the productivity at each history and the labor disutility

[27] In fact, in the dataset some permanent workers cross the threshold between age groups.

[28] 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.

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

parameter Γ. The productivities are pinned down with the first-order condition of labor supply

$$\theta = y(\theta)\sqrt{\Gamma\frac{1 - T(y(\theta))/y(\theta)}{1 - T'(y(\theta))}}, \tag{10}$$

where $T(y)$ is the actual tax Italian tax schedule.[29] Without the loss of generality Γ is set to 1 - by the first order condition (10) varying Γ would simply rescale all the productivities. The discount factor β is equal to 0.5, corresponding to the period of 17 years.

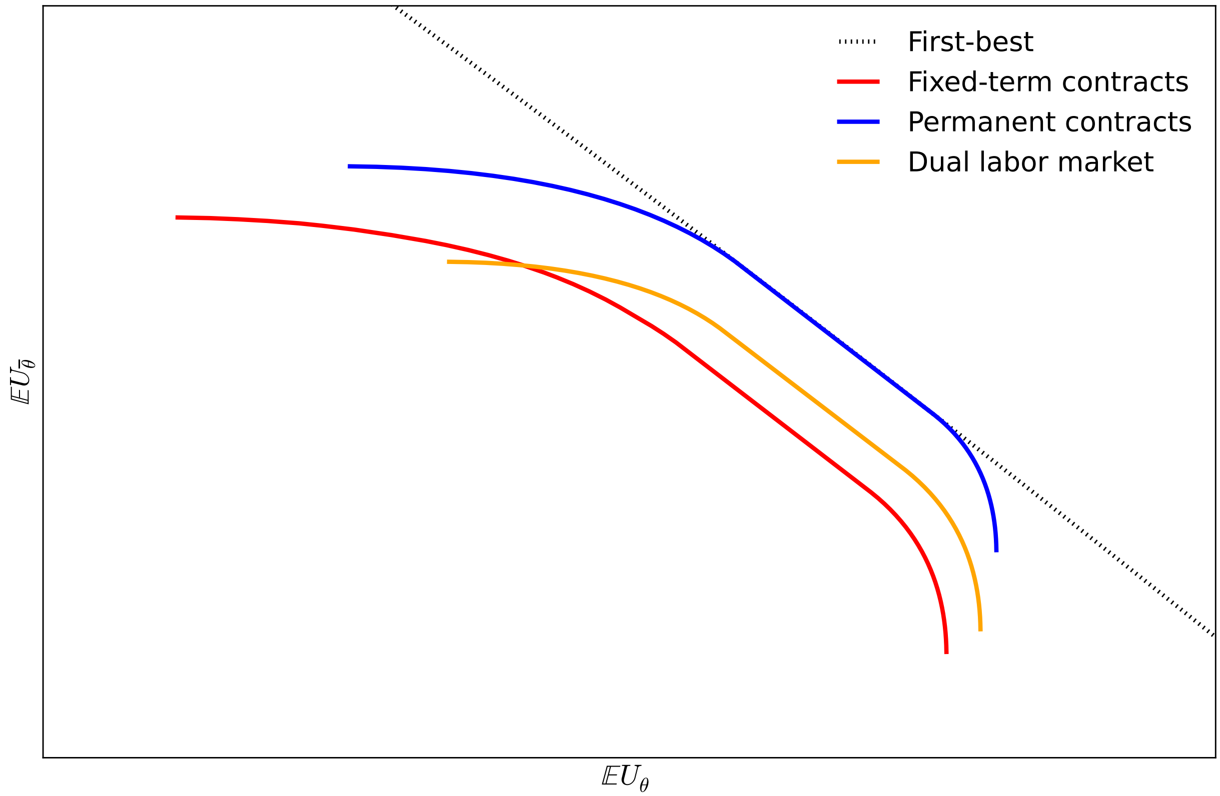
7.2 Pareto Frontiers

Figure 1 shows 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.[30] 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 regime the government raises the same net tax revenue as the actual Italian tax schedule.

[29]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.

[30]I do not plot the Pareto frontier of the fourth configuration, in which the initially low productivity type receives permanent contract and the initially high type receives fixed-term contract, as it is already implicitly incorporated in the 'Permanent contracts' frontier - see Lemma A.3.

Figure 1: Pareto frontiers in the calibrated economy



Whenever the worker is employed on 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, the Pareto frontier of any regime in which some workers are employed on a fixed-term basis is bounded away from the first-best Pareto frontier. On the other hand, permanent contracts allow for coexistence of full insurance and efficient labor supply if the redistribution between initial types is limited. The permanent contracts frontier coincides with the first-best when the social preferences are not strongly redistributive.

By Theorem 3 we know that the regime with only fixed-term contracts is Pareto dominated by some regime in which at least some workers receive permanent contracts. Figure 1 shows that for the calibrated parameter values it is always optimal to assign permanent contracts to all workers. Furthermore, in any constrained efficient allocation all workers enjoy full consumption insurance.[31] The dual labor market regime improves upon the fixed-term regime when the planner cares predominantly about the initially less productive workers, but the gains from assigning permanent contract to both types are even greater. The welfare gap between the permanent contracts and the dual labor market regimes is 0.6% in consumption equivalent terms.

[31] For some parameter values when the planner wants to redistribute from the bottom to the top, a dual labor market in which initial high productivity workers have fixed-term contract is optimal. In such case the permanent contracts allocation would not involve full consumption insurance of the high productivity workers.

Table 5: Optimal allocations for different social welfare functions

Social preferences:	**utilitarian**	**libertarian**	**Rawlsian**	**anti-Rawlsian**
Objective	$\mathbb{E}U$	$\sum_{\theta\in\Theta_1} \lambda^{lf}(\theta)\,\mu(\theta)\,\mathbb{E}U_\theta$	$\min_{\theta\in\Theta_1}\mathbb{E}U_\theta$	$\max_{\theta\in\Theta_1}\mathbb{E}U_\theta$
Consumption insurance	full	full	full	full
Allocation of labor	efficient	efficient	distorted	distorted
Welfare (cons. equiv.)				
laissez-faire = 100%	104.3%	104.1%	108.3%	107.2%
NDPF = 100%	101.5%	101.5%	101%	101.2%
Relative gain from permanent contracts	53.3%	59.7%	12.9%	20.3%

Note: The libertarian planner maximizes individual utility subject to no redistribution between initial types. Taxation plays only an insurance role and workers would voluntarily decide to participate in a public insurance scheme. The libertarian welfare weights λ^{lf} are welfare weights under which in the first-best there is no redistribution. The anti-Rawlsian planner is the opposite of the Rawlsian planner and maximizes the utility of the most well-off type.

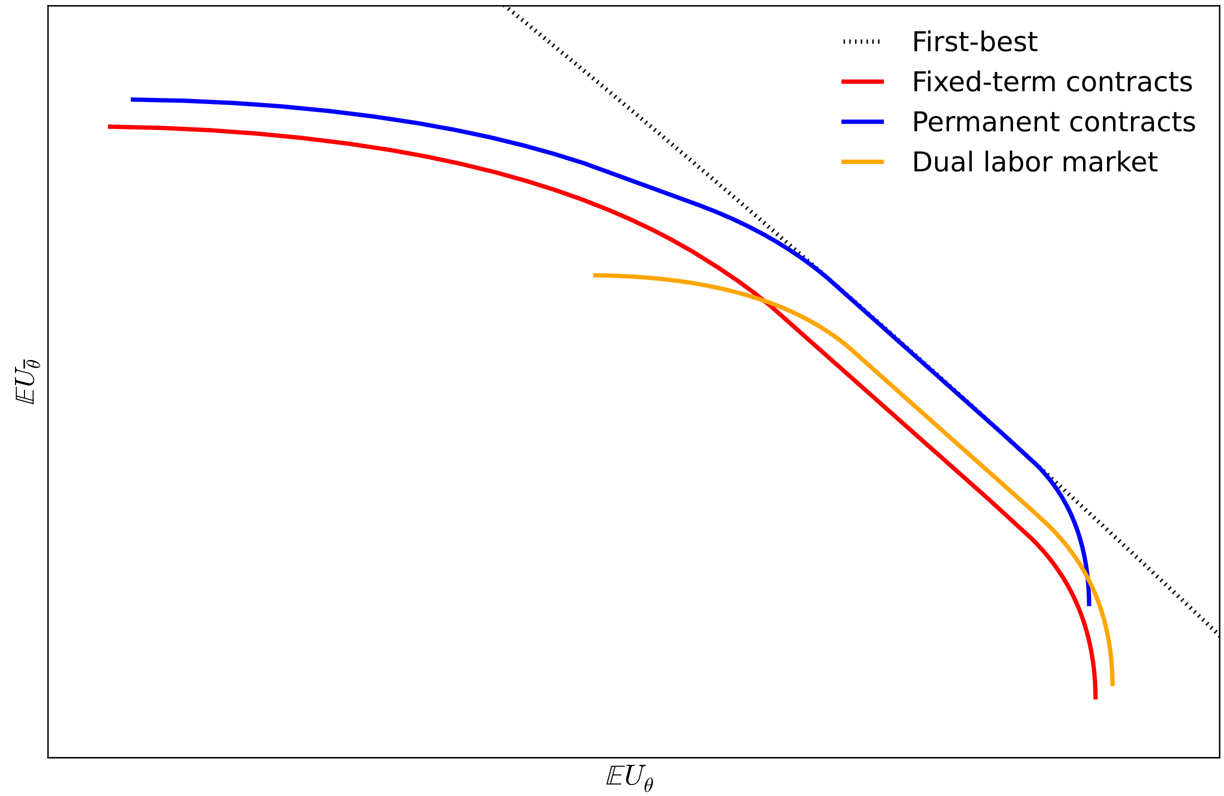
Table 5 presents the optimal allocations under different assumptions about the social preferences. Interestingly, the utilitarian planner does not introduce labor distortions and chooses the allocation that lays on the first-best Pareto frontier.[32] Labor distortions augment redistribution and are used only by the two most redistributive social planners: Rawlsian and anti-Rawlsian, who care only about one initial type of workers. The middle rows of Table 5 compare the constrained efficient allocations with the laissez-faire (fixed-term contracts, no public insurance, uniform lump-sum taxation to cover the government expenditures) 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 social planners (utilitarian and libertarian), the NDPF captures less than two-thirds of the gains from constrained efficient allocation. It means that the simpler tax systems that encourages firms' insurance improve upon the complicated tax systems prescribed by NDPF by more than 50% in terms of welfare (the last row of Table 5). The relative welfare gain from using permanent contracts in comparison to fixed-term contracts is smaller for social preferences focused on redistribution.

Consider the alternative economy in which the initial differences between two types are greater than in Italy: suppose that the initial productivity of less productive type is lower by 10%. The Pareto frontiers of different regimes are shown on Figure 2. Now the Rawlsian planner implement dual labor markets. The greater difference in initial productivities means that the deviating high type gains more by smoothing labor, which is not possible when the initial low type has fixed-term contract. The welfare gain of the dual labor market over the permanent contract regime is large: 1.8% in consumption equivalent terms.[33] Note that in this case the Rawlsian planner prefers even the NDPF allocation to the permanent contracts allocation. The dual labor market becomes optimal also when we increase the population share of initial high productivity type in comparison to the calibrated case.

[32] Although the constrained efficient allocation chosen by the utilitarian planner lays on the first-best Pareto frontier, it does not coincide with the first-best utilitarian optimum. The planner 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.

[33] The dual labor market and the permanent contract regime yield roughly the same welfare when the initial low productivity is lower by 4% in comparison to the calibrated value.

Figure 2: Pareto frontiers in the alternative economy with increased initial differences



8 Conclusions

Firms are the natural insurers of their employees. First, competition in the labor market gives firms strong incentives to shelter workers from risk. Second, companies arguably have the best knowledge of their workers' productivity (with an exception of workers themselves). The insurance role of firms should be acknowledged by the optimal tax theory, as it can resolve its shortcoming: the excessive complexity. In this paper I show that incorporating firms into the dynamic taxation framework leads to a fully optimal tax system that is remarkably simple. The government optimally outsources insurance to firms by promoting permanent employment contracts. The remaining role for the government is redistribution, which can be conducted with simple tax instruments. In a calibrated model of Italy all constrained efficient allocations can be implemented with a time-invariant tax schedule that depends only on current consumption expenditure.

Empowering the private sector to insure workers comes at a cost. Firms insure their workers by shifting income from the times of high to the times of low productivity. However, this intertemporal reallocation can be used to avoid taxes in the following way: a productive worker shifts the income to the future and collects income support today. A redistributive government can limit such a behavior without reducing the generosity of transfers by promoting fixed-term contracts at low earnings levels. This redistributive argument provides a novel perspective on dual labor markets in which permanent and fixed-term contracts coexist.

The analysis could be extended in several directions. The focus of this paper is on workers' heterogeneity. Introduction of heterogenous production opportunities for firms would allow for analysis of temporary jobs, which in fact are the primary reason for existence of fixed-term contract. Limited commitment of workers may be muted by search and mobility costs. Careful treatment of employees' outside option is vital to understanding insurance within firm. Finally, it is worth examining how the trade-off between redistribution and insurance influences the optimal regulation of other insurance markets.

References

Abraham, A., S. Koehne, and N. Pavoni (2014): "Optimal income taxation with asset accumulation," .

Albanesi, S., and C. Sleet (2006): "Dynamic optimal taxation with private information," *The Review of Economic Studies*, 73(1), 1–30.

Attanasio, O., and J.-V. Rios-Rull (2000): "Consumption smoothing in island economies: Can public insurance reduce welfare?," *European Economic Review*, 44(7), 1225–1258.

Bentolila, S., P. Cahuc, J. J. Dolado, and T. Le Barbanchon (2012): "Two-Tier Labour Markets in the Great Recession: France Versus Spain*," *The Economic Journal*, 122(562), F155–F187.

Best, M. C. (2014): "Salary Misreporting and the Role of Firms in Workers Responses to Taxes: Evidence From Pakistan," Discussion paper, Working Paper, London School of Economics.

Cabrales, A., J. J. Dolado, and R. Mora (2014): "Dual Labour Markets and (Lack of) On-the-Job Training: PIAAC Evidence from Spain and Other EU Countries," .

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

Cole, H. L., and N. R. Kocherlakota (2001): "Efficient allocations with hidden income and hidden storage," *The Review of Economic Studies*, 68(3), 523–542.

Conesa, J. C., S. Kitao, and D. Krueger (2009): "Taxing Capital? Not a Bad Idea After All!," *American Economic Review*, 99(1), 25–48.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977): "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the royal statistical society. Series B (methodological)*, pp. 1–38.

Doligalski, P., and L. Rojas (2015): "Optimal Redistribution with a Shadow Economy," .

Eichhorst, W. (2014): "Fixed-term contracts," *IZA World of Labor*.

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

Findeisen, S., and D. Sachs (2015): "Redistribution and Insurance with Simple Tax Instruments," .

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

García-Pérez, J. I., I. Marinescu, and J. V. Castello (2014): "Can Fixed-Term Contracts Put Low Skilled Youth on a Better Career Path? Evidence from Spain," .

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal indirect and capital taxation," *Review of Economic studies*, pp. 569–587.

Golosov, M., M. Troshkin, and A. Tsyvinski (2015): "Redistribution and Social Insurance," .

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

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

Golosov, M., A. Tsyvinski, and I. Werning (2007): "New dynamic public finance: a user's guide," in *NBER Macroeconomics Annual 2006, Volume 21*, pp. 317–388. MIT Press.

Guiso, L., L. Pistaferri, and F. Schivardi (2005): "Insurance within the Firm," *Journal of Political Economy*, 113(5).

——— (2013): "Credit within the Firm," *The Review of Economic Studies*, 80(1), 211–247.

Guvenen, F., F. Karahan, S. Ozkan, and J. Song (2015): "What do data on millions of US workers reveal about life-cycle earnings risk?," Discussion paper, National Bureau of Economic Research.

Handel, B. R., I. Hendel, and M. D. Whinston (2013): "Equilibria in Health Exchanges: Adverse Selection vs. Reclassification Risk," Discussion paper, National Bureau of Economic Research.

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

Hendel, I., and A. Lizzeri (2003): "THE ROLE OF COMMITMENT IN DYNAMIC CONTRACTS: EVIDENCE FROM LIFE INSURANCE.," *Quarterly Journal of Economics*, 118(1).

Huggett, M., G. Ventura, and A. Yaron (2011): "Sources of Lifetime Inequality," *The American Economic Review*, 101(7), 2923.

Ichino, A., M. Polo, and E. Rettore (2003): "Are judges biased by labor market conditions?," *European Economic Review*, 47(5), 913–944.

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

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

Kosior, A., M. Rubaszek, and K. Wierus (2015): "On the importance of the dual labour market for a country within a monetary union," .

Krueger, D., and F. Perri (2011): "Public versus private risk sharing," *Journal of Economic Theory*, 146(3), 920–956.

Krueger, D., and H. Uhlig (2006): "Competitive risk sharing contracts with one-sided commitment," *Journal of Monetary Economics*, 53(7), 1661–1691.

Lagakos, D., and G. L. Ordonez (2011): "Which workers get insurance within the firm?," *Journal of Monetary Economics*, 58(6), 632–645.

Marin, J.-M., K. Mengersen, and C. P. Robert (2005): "Bayesian modelling and inference on mixtures of distributions," *Handbook of statistics*, 25(16), 459–507.

Milgrom, P., and C. Shannon (1994): "Monotone comparative statics," *Econometrica: Journal of the Econometric Society*, pp. 157–180.

Mirrlees, J. A. (1971): "An exploration in the theory of optimum income taxation," *The review of economic studies*, pp. 175–208.

Parsons, D. O. (2002): "Private 'Unemployment'Insurance: Asymmetric Information and Severance Pay Design," Discussion paper, mimeo, George Washington University, Washington.

Saez, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68(1), 205–29.

Stantcheva, S. (2014): "Optimal Income Taxation with Adverse Selection in the Labour Market," *The Review of Economic Studies*, p. rdu005.

Tealdi, C. (2011): "Typical and atypical employment contracts: the case of Italy," .

Thomas, J., and T. Worrall (1988): "Self-enforcing wage contracts," *The Review of Economic Studies*, 55(4), 541–554.

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

Weinzierl, M. (2011): "The surprising power of age-dependent taxes," *The Review of Economic Studies*, 78(4), 1490–1518.

White, H. (1980): "A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity," *Econometrica: Journal of the Econometric Society*, pp. 817–838.

A Proofs and auxiliary lemmas

A.1 Section 2

Proof of Lemma 1. Suppose that there is an equilibrium contract (σ, n) which at some history $h \in \mathcal{H}$ yields positive profits. A competitor could profitably steal the worker by offering contract (σ', n'), where $\sigma' = \sigma$, $n'(h) = n(h) - \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}\pi_h(c \circ r, n)/2\theta(h)$ and $n'(s) = n(s)$ for any other history $s \neq h$. This contract yields half of the profits of the original contract (r, n) and would be preferred by the worker, as it involve less labor supply. Suppose instead that profits are negative and lower than $f(r(h))$. Then the firm would prefer to terminate the contract and incur firing cost rather than keep the worker. Hence, all equilibrium contracts satisfy $\forall_{h\in\mathcal{H}} f(r(h)) \leq \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}\pi_h(c \circ r, n) \leq 0$. Moreover, the firm cannot expect losses initially, as it could offer a contract that yields zero profits by equalizing worker's output to worker's income at each history. Hence, $\forall_{\theta\in\Theta_1} \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}\pi_\theta(c \circ r, n) = 0$.

Suppose there is an equilibrium contract (σ, n) which does not belong to $\mathcal{E}(c, y, f)$. It means there is another contract (σ', n') which yields strictly greater expected utility to the worker subject to the zero profit condition and limited commitment constraints. Define $\sigma'' = \sigma'$ and $\forall_{h\in\mathcal{H}} n''(h) = n'(h) + \epsilon$, where $\epsilon > 0$. For epsilon sufficiently small (σ'', n'') yields greater expected utility than (σ, n) and positive profits. Hence, (σ, n) cannot be an equilibrium contract.

Suppose that there is a contract $(\sigma, n) \in \mathcal{E}(c, y, f)$ that is not an equilibrium contract. It means that there is another contract (σ', n') that yields positive profits and the expected utility greater than (σ, n). This in turn implies that there is yet another contract which yields zero profits at each history and the expected utility greater than (σ', n'). It contradicts the fact that $(\sigma, n) \in \mathcal{E}(c, y, f)$. □

A.2 Section 3

Proof of Theorem 1. I will show that the 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: $\forall_{h\in\mathcal{H}_{\bar{t}}(\theta_1)} \sum_{t=1}^{\bar{t}} R^{1-t} y(h^t) = Y(h_1)$.

Take any outcomes (c, y) which are incentive-compatible (i.e. σ^* is the equilibrium reporting strategy) and do not involve full consumption insurance. For each initial type $h_1 \in \times_1$ find a full-insurance consumption level $\bar{c}(h_1)$ with equality $\sum_{t=1}^{\bar{t}} \beta^{t-1} u(\bar{c}(h_1)) = \sum_{s\in\mathcal{H}(h_1)} \beta^{1-|s|} \mu(s \mid h_1) u(c(h))$. Set $\bar{y}(h)$, where $h \in \mathcal{H}(\theta)$, to the average of histories of this length: $\sum_{s\in\mathcal{H}_{|h|}(\theta)} \mu(s \mid \theta) y(s)$. This way both the lifetime consumption and labor income are deterministic functions of the initial type report. The worker has the same lifetime output as before (as it is equal to the lifetime income). Futhermore, the worker receives more consumption insurance, which allows the planner to save some resources. Before we prove that truthtelling is the equilibrium strategy given the new outcomes, lets define a useful reporting strategy.

Definition 8. The "statistical mimicking" strategy $\sigma^{stat}_{\theta,\theta'}$ is a reporting strategy such that

- for all $r \in \mathcal{R}$ such that $\exists_{h\in\mathcal{H}\setminus\mathcal{H}(\theta)} r(h) \neq h$ implies that $\sigma_{\theta,\theta'}(r) = 0$.

- $\forall_{h\in\mathcal{H}(\theta')}\mu\left(h\mid\theta'\right)=\sum_{s\in\mathcal{H}_{|h|}(\theta)}\mu\left(s\mid\theta\right)\sum_{r\in\mathcal{R}}\mathbb{I}_{r(s)=h}\sigma\left(r\right)$, where $\mathbb{I}$ is an indicator function.

The first point means that apart from the history following the initial type θ, $\sigma_{\theta,\theta'}^{stat}$ is truthful. The second point means that along the history following the initial type θ, $\sigma_{\theta,\theta'}^{stat}$ generates the distribution of type reports consistent with truthful reporting of initial type θ'.

In order to check that the truthful reporting is the equilibrium strategy given the new outcomes, note that the initial type report uniquely determines the payoff of the worker. Consider a reporting strategy $\bar{r}$ such that $\bar{r}\left(\theta_1\right)=\theta_2$ (involves type θ_1 mimicking type θ_2) and $\bar{r}\left(\theta\right)=\theta$ for $\theta\in\Theta_1\backslash\theta_1$ (is truthful elsewhere). The utility from following this strategy given outcomes $\left(\bar{c},\bar{y}\right)$ is equal to the utility from following the strategy $\sigma_{\theta_1,\theta_2}^{stat}$ with outcomes $\left(c,y\right)$. To see this, note that the utility from consumption and the expected lifetime income generated by $\sigma_{\theta_1,\theta_2}^{stat}$ are identical to those generated by truthful reporting of type θ_2. Since truthful reporting was a equilibrium reporting strategy given outcomes (c,y), the truthful reporting is an equilibrium strategy also with outcomes $\left(\bar{c},\bar{y}\right)$. Hence, the planner can get free resources without hurting any worker by implementing the full consumption insurance.

Now I will show that the full consumption insurance is incentive compatible (i.e. implies a truthful equilibrium reporting) is and only is if and only if the lifetime labor income along any history depends only on the initial type:

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

Denote by $\underline{h}$ $(\overline{h})$ a history originating from θ that yields the lowest (highest) lifetime income:

$$\underline{h}\in\arg\min_{h\in\mathcal{H}_{\bar{t}}(\theta)}\sum_{t=1}^{\bar{t}}R^{1-t}y\left(h^t\right),\ \overline{h}\in\arg\min_{h\in\mathcal{H}_{\bar{t}}(\theta)}\sum_{t=1}^{\bar{t}}R^{1-t}y\left(h^t\right).$$

Suppose that the lifetime income is strictly greater along $\overline{h}$ in comparison to $\underline{h}$. Because of full insurance, consumption allocation is identical along $\underline{h}$ and $\overline{h}$. By the zero profit condition, the worker can reduce the average level of labor supply by always reporting $\underline{h}$. Hence, the equilibrium labor contract involves reporting $\underline{h}$ whenever $\overline{h}$ happens. By contradiction, it means that whenever there is a full consumption insurance, the lifetime income has to be a deterministic function of the initial type. 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 by the zero profit condition means that the life-time income at the history s is greater than at the history h. □

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

Proof of Lemma A.1. I will show that under Assumption 1 $\frac{\partial V_{\theta_1}(C,Y)}{\partial Y}$ is increasing with θ. Then I will show that it implies that $V_\theta(C,Y)$ has the single crossing property, as in Milgrom and Shannon (1994).

First note that $\frac{\partial V_{\theta_1}(C,Y)}{\partial Y} = -\frac{v'(n(h))}{\theta(h)}$ for any $h \in \mathcal{H}(\theta_1)$. Suppose that $\frac{v'(n(h))}{\theta(h)}$ does not decrease with θ_1. It means that for all s such that $s \in \times_{i=2}^{k}\Theta_i$, $k \leq \bar{t}$, we have $n(\theta_1, s) \leq n(\theta_2, s)$ for $\theta_2 > \theta_1$. Then the following inequalities are implied by the weak FOSD (Assumption 1)

$$\forall_{t\in\{1,\ldots,\bar{t}\}} \sum_{s\in\mathcal{H}_t(\theta_1)} \mu(s \mid \theta_1)\theta(s)n(s) \leq \sum_{s\in\mathcal{H}_t(\theta_2)} \mu(s \mid \theta_2)\theta(s)n(s).$$

At least one, initial inequality ($t = 1$) is always strict, so the initial type θ_2 produces strictly more in expected value than θ_1. It leads to a contradiction, as the initial zero-profit condition is slack, which cannot happen in equilibrium.

Milgrom and Shannon (1994) strict single crossing condition means that

$$V_{\theta_1}(C_2, Y_2) \geq V_{\theta_1}(C_1, Y_1) \implies V_{\theta_2}(C_2, Y_2) > V_{\theta_2}(C_1, Y_1),$$

where $\theta_2 > \theta_1$ and $Y_2 > Y_1$ (which naturally implies that $C_2 > C_1$). Given the additive separability between consumption and labor, this statement is true if the following holds for any C:

$$V_{\theta_1}(C, Y_1) - V_{\theta_1}(C, Y_2) > V_{\theta_2}(C, Y_1) - V_{\theta_2}(C, Y_2).$$

Since V_θ is monotone in Y, we can write this inequality as

$$\int_{Y_1}^{Y_2} \frac{dV_{\theta_1}(C,Y)}{dY} dY < \int_{Y_1}^{Y_2} \frac{dV_{\theta_2}(C,Y)}{dY} dY,$$

which is true, since, as it was shown before, $\frac{dV_\theta(C,Y)}{dY}$ is increasing in θ. $\square$

Proof of Proposition 1 . By Theorem 1 the planner's problem under full commitment can be expressed 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. The single crossing holds by Assumption 1 and Lemma A.1. 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 (6) 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 the zero profit condition to get

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

By the implicit function theorem we have

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

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

Hence, we have

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

where $\zeta^c\left(h\right)$ 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

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

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

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

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

where $\xi\left(h\right)$ is the wealth effect at the history h. The lifetime wealth effect is the average wealth effect across all histories. Now what remains to be done is to plug the derived elasticities in the Saez (2001) formula. □

A.3 Section 4

Lemma A.2. *If workers cannot commit, the payoff from any mixed reporting strategy is not greater than the payoff from some pure reporting strategy.*

Proof of Lemma A.2 . Suppose that there is some initial type θ and a mixed reporting strategy σ that randomizes between two pure reporting strategies r_1 and r_2 with probabilities such that

$$\sum_{r\in\mathcal{R}}\sigma(r)\mathbb{E}U_\theta\left(c\circ r,\tilde{n}\right)>\sigma(r_1)\mathbb{E}U_\theta\left(c\circ r_1,n_1\right)+\sigma(r_2)\mathbb{E}U_\theta\left(c\circ r_2,n_2\right),$$

where n_1, n_2 and $\tilde{n}$ are the labor supply allocations corresponding to the three reporting strategies. Since the consumption allocation is fixed by the mechanism, we can subtract it from both sides. The only source of utility gain could only be a more efficient allocation of labor

$$\sum_{h\in\mathcal{H}(\theta)}\beta^{1-|h|}\mu\left(h\right)v\left(\tilde{n}\left(h\right)\right)>\sigma(r_1)\sum_{h\in\mathcal{H}(\theta)}\beta^{1-|h|}\mu\left(h\right)v\left(n_1\left(h\right)\right)+\sigma(r_2)\sum_{h\in\mathcal{H}(\theta)}\beta^{1-|h|}\bar{\mu}\left(h\right)v\left(n_2\left(h\right)\right). \tag{11}$$

The inequality requires that the expected lifetime income under r_1 and r_2 is different, since otherwise otherwise n_1, n_2 and $\tilde{n}$ would be identical.[34] Without the loss of generality suppose that $\sum R^{1-|h|}\mu\left(h\right)y\left(r_1\left(h\right)\right)<\sum R^{1-|h|}\mu\left(h\right)y\left(r_2\left(h\right)\right)$. As the expected lifetime labor income implied by the labor policy on the left-hand and the right-hand side of (11) are equal, the only source of lower disutility from labor can be more efficient allocation of labor. As n_1 and n_2 are optimized conditional on each reporting strategy, it has to be the case that $\tilde{n}$ smooths labor supply across the two pure reporting strategies. Such smoothing requires that the agent that has drawn r_1 works a bit more than in n_1 and the agent that has drawn r_2 works a bit less than n_2. However, it means that the agent that has drawn r_1 would benefit from taking a pure labor contract $\left(r_1,n_1\right)$. Hence, the randomized strategy $(\tilde{n},\sigma)$ is not an equilibrium contract. Since it has to be the case that

$$\sum_{r\in\mathcal{R}}\sigma(r)\mathbb{E}U_\theta\left(c\circ r,\tilde{n}\right)\leq\sigma(r_1)\mathbb{E}U_\theta\left(c\circ r_1,n_1\right)+\sigma(r_2)\mathbb{E}U_\theta\left(c\circ r_2,n_2\right),,$$

the mixed reporting strategy cannot bring more utility than at least one of the two pure reporting strategies. □

Proof of Lemma 2. With fixed-term contract and given a reporting strategy r, the limited commitment constraints mean that $\forall_{h\in\mathcal{H}}\mathbb{E}U_h(c\circ r,n)=0$. I will show by induction that it implies that income and output coincide at each history. Zero profit conditions at the histories of length $\bar{t}$ imply $\forall_{h\in\mathcal{H}_{\bar{t}}}y\left(r(h)\right)=\theta\left(h\right)n\left(h\right)$. Consider history of length t and suppose that for all histories of length greater than t labor income equals output. Then $\forall_{h\in\mathcal{H}_t}\mathbb{E}\pi_h(c\circ r,n)=\theta\left(h\right)n\left(h\right)-y\left(r(h)\right)$, which is equal to zero by the zero profit condition. □

Proof of Theorem 2. Take any allocation under full commitment on the labor market: (c^{FC},y^{FC},n^{FC}), where $(\sigma^*,n^{FC})\in\mathcal{E}^{FC}(c^{FC},y^{FC})$. We will find a new allocation of labor income y such that $(\sigma^*,n^{FC})\in\mathcal{E}(c^{FC},y,f^P)$, where only permanent contracts are used: $\forall_{h\in\mathcal{H}}f^P(h)=\bar{f}$. For any

[34] Conditional on reporting strategy, the optimal labor supply is depends only on the expected lifetime income.

non-initial history $s \in \mathcal{H} \backslash \mathcal{H}_1$ set $y(s) = \max_{s' \in \mathcal{H}_{|s|}(s^{-1})} \theta(s')n(s')$. The limited commitment constraint holds at $s : \mathbb{E}\pi_s(y, n) \leq 0$, as the firm pays the worker at least his output. Then modify the initial labor income such that the zero profit condition holds:

$$\forall_{h \in \mathcal{H}_1} y(h) = y^{FC}(h) - \sum_{s \in \mathcal{H}(h) \backslash \{h\}} R^{1-|s|} \mu(s \mid h)(y(s) - y^{FC}(s)).$$

As the expected lifetime income of each initial type is unchanged, the zero profit condition holds. Hence, (σ^*, n^{FC}) belongs to the constraint set of the maximization problem that defines $\mathcal{E}(c^{FC}, y, f^P)$. To see that in the constraint set there is no better contract for the worker, note that if there was, then (σ^*, n^{FC}) would not be an equilibrium contract under full commitment. □

Lemma A.3. *Suppose that (i) there are two time periods, (ii) there are two initial types: $\overline{\theta} > \underline{\theta}$, (iii) second period productivity is independent of the initial type. Consider allocation (c, y, n) implemented with mechanism (c, y, f), where $f(\overline{\theta}) = 0$. The planner can implement the same allocation with a mechanism (c, y, f'), where $f'(\overline{\theta}) = \bar{f}$ and $f'(\underline{\theta}) = f(\underline{\theta})$.*

Proof of Lemma A.3. Since there is no redistribution from the high to the low type, the labor supply of $\underline{\theta}$ will be undistorted, while the labor supply of $\overline{\theta}$ can be either undistorted or distorted upwards. Consider the mechanism with the contract assignment f'. When $\underline{\theta}$ mimics the other type, he could in principle use labor smoothing. Since $\underline{\theta}$ is comparatively less productive initially, the beneficial labor smoothing would require producing more in the second period. However, the limited commitment constraints do not allow to increase the second period output above income. Hence, the low type cannot gain from labor smoothing and the utility from deviation is the same regardless of type of contract of $\overline{\theta}$. The truthful revelation is still an equilibrium contract under f'. □

Proof of Theorem 3. Suppose that some top taxpayer θ has fixed-term contract. Assign this type permanent contract and fully smooth consumption, while keeping the net present value of consumption and income constant. This type is better-off. If any other type is tempted to mimic θ, assign them the same outcomes as for θ. They are weakly better-off as well. Furthermore, the planner has more resources, as the mimicking types are now contributing more to the resource constraint. □

Proof of Lemma 3. $\Delta^{insurance}$ is non-positive, since both c_1 and c^P have the same net present value and are chosen to maximize the utility of $\underline{\theta}$, but c_1 has to satisfy additional incentive-compatibility constraints. $\Delta^{redistribution}$ is non-negative, because the incentive constraints between initial types are always (weakly) relaxed by introducing fixed-term contracts. $\Delta^{efficiency}$ is non-negative, since otherwise the planner in the dual labor market case would chose the allocation (c_2, y^P, f^D), which implies that $\Delta^{efficiency} = 0$. □

Proof of Proposition 2. Suppose that the planner maximizes the utility of $\underline{\theta}$. Denote by $C(\theta)$ and $Y(\theta)$ the present value of consumption and labor income of each initial type. Suppose first that $\underline{\theta}$

has a permanent contract. The binding incentive constraint is

$$C(\overline{\theta}) - \sum_{s \in \mathcal{H}(\overline{\theta})} \beta^{|s|-1} \mu(s \mid \overline{\theta}) v(n_0(s)) \geq C(\underline{\theta}) - \sum_{s \in \mathcal{H}(\overline{\theta})} \beta^{|s|-1} \mu(s \mid \overline{\theta}) v(\tilde{n}_0(s)), \tag{12}$$

where $\tilde{n}_0$ produces $Y(\underline{\theta})$ and satisfies the labor smoothing condition (6). Suppose instead that $\underline{\theta}$ has a fixed-term contract. Since types are independently distributed, the biding constraint is

$$C(\overline{\theta}) - \sum_{s \in \mathcal{H}(\overline{\theta})} \beta^{|s|-1} \mu(s \mid \overline{\theta}) v(n_0(s)) \geq C(\underline{\theta}) - \sum_{s \in \mathcal{H}(\overline{\theta})} \beta^{|s|-1} \mu(s \mid \overline{\theta}) v(n'(s)), \tag{13}$$

where $n'(s) = \frac{y(\underline{\theta}, s_2, .., s_s)}{\theta(\overline{\theta}, s_2, .., s_s)}$. I will show that the right-hand side of (12) is strictly greater than of (13). Note that both $\tilde{n}_0$ and n' produce the same output $Y(\underline{\theta})$. Note that n satisfies

$$\forall_{s \in \mathcal{H}(\underline{\theta})} \frac{v'(n(\underline{\theta}))}{\underline{\theta}} = \frac{v'(n(s))}{\theta(s)}.$$

It implies that

$$\forall_{s \in \mathcal{H}(\overline{\theta})} \frac{v'(n'(\overline{\theta}))}{\overline{\theta}} < \frac{v'(n'(s))}{\theta(s)}.$$

The initial type $\overline{\theta}$ could reduce the disutility from labor by producing more in the initial period., which is consistent with limited commitment constraints. It means that $n'(\overline{\theta}) < \tilde{n}_0(\overline{\theta})$ and hence the right-hand side of (12) is strictly greater than of (13). Hence, the incentive constraint is relaxed when $\underline{\theta}$ receives fixed-term contract and $\Delta^{redistribution} > 0$.

□

Proof of Proposition 3. Under permanent contract, consumption of type $\underline{\theta}$ in each period is constant. If we give this type 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 is independently distributed, introducing distortions after the initial period does not help with redistribution and is driven purely by insurance motives. 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

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

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

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

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. □

Proof of Lemma 4. It follows from Lemma 1. □

Proof of Theorem 4. First I will show that there exists a history dependent asset policy a such that, given the tax T, the allocation together with a belongs to $\tilde{\mathbb{E}}$. Define $a(h) = \sum_{t=1}^{|h|} R^{1-|h|}(y(h^t) - \bar{y}(h_1))$. Then at each history consumption expenditure is equal $\bar{y}(h_1)$. Hence at each history $T_x(x(h)) = \bar{T}(\theta)$ and the right-hand side of the budget constraint equals $c(h)$.

Suppose that $\bar{T}_x$ is convex. Take some initial type θ. He may deviate either to a constant, but different level of consumption expenditures, or he may introduce some volatility to consumption expenditures. The first deviation is taken care of by the equilibrium constraint from the corresponding direct mechanism as well as a punitively high tax whenever worker deviates to a level that do not correspond to consumption expenditures of any other tax. The introduction of volatility in consumption expenditures with the expected value $\bar{x}$ means, due to convexity of the tax system, that the expected tax paid is not lower than $T_x(\bar{x})$. Since the labor supply allocation is the same both cases (labor allocation depends on expected lifetime income, which is equal to the expected lifetime expenditure), the utility from introducing volatility is bounded above by a utility of deviation to the constant consumption expenditure $\bar{x}$ (which was taken care of above). Note that deviations to fixed-term contract imply volatile consumption expenditures and are not tempting by the same argument.

When $\bar{T}_x$ is not convex, we can add an auxiliary correction term $\alpha(x_t - x_0)^2$ which punishes volatile consumption expenditures. The parameter α should be high enough such $\bar{T}_x\left(x_t\right) + \alpha\left(x_t - x_0\right)^2$ is convex. □

Proof of Proposition 4. I will show that the inverse Euler equation holds. The proof follows the intuition from Golosov, Kocherlakota, and Tsyvinski (2003). Take any allocation (c, y, n) implemented by some direct mechanism (c, y, f). Take some history $h \in \mathcal{H} \backslash \mathcal{H}_{\bar{t}}$ and consider a small perturbation δ such that

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

and c' equal c elsewhere. As the utility from any reporting strategy is unchanged, truthful revelation still holds in equilibrium. In the optimum such perturbation cannot yield free resources

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

which implies that the inverse Euler equation holds. To see how the inverse Euler equation together with volatile consumption implies a savings distortion and capital tax, see for instance Golosov, Kocherlakota, and Tsyvinski (2003). □

B Auxiliary estimates

Table 6: The regression of $\log\left(\frac{y_{ijt}}{y_{ijt-1}}\right)$ for permanent contracts.

variable	**coef**	**std err**	**t**	**p-value**
const	0.0285***	0.001	27.766	0.000
male	0.0072***	0.001	11.940	0.000
tenure	-0.0015***	4.97e-05	-30.120	0.000
d1999	-0.0094***	0.001	-11.243	0.000
d2000	-0.0073***	0.001	-8.857	0.000
d2001	-0.0056***	0.001	-6.812	0.000
d2002	-0.0095***	0.001	-11.697	0.000
white_colar	0.0169***	0.001	28.766	0.000
cadre	0.0106***	0.001	7.288	0.000
manager	-0.0391***	0.002	-18.043	0.000
North-East	0.0008	0.001	1.274	0.202
Center	-0.0036***	0.001	-4.959	0.000
South	-0.0056***	0.001	-6.717	0.000
Islands	-0.0055***	0.001	-4.616	0.000
young_firm	0.0011*	0.001	1.660	0.097
agriculture	-0.0096**	0.005	-2.093	0.036
heavy_industry	-0.0072***	0.002	-3.892	0.000
manufacturing	-0.0038***	0.001	-6.149	0.000
construction	-0.0174***	0.001	-16.560	0.000
services1	0.0002	0.001	0.224	0.823

Table 7: The regression of $\log\left(\frac{y_{ijt}}{y_{ijt-1}}\right)$ for fixed-term contracts.

variable	**coef**	**std err**	**t**	**p-value**
const	0.0388***	0.010	3.753	0.000
male	0.0031	0.006	0.541	0.588
tenure	-0.0037***	0.001	-3.464	0.001
d1999	-0.0042	0.010	-0.421	0.674
d2000	0.0053	0.009	0.558	0.577
d2001	-6.927e-05	0.009	-0.007	0.994
d2002	-5.987e-05	0.009	-0.007	0.994
white_colar	0.0323***	0.006	5.358	0.000
cadre	0.0103	0.034	0.302	0.763
manager	-0.0192	0.037	-0.517	0.605
North-East	-0.0111*	0.007	-1.643	0.100
Center	-0.0165**	0.007	-2.293	0.022
South	-0.0031	0.009	-0.341	0.733
Islands	-0.0213*	0.011	-1.902	0.057
young_firm	-0.0047	0.007	-0.698	0.485
agriculture	-0.0392	0.041	-0.955	0.339
heavy_industry	-0.0006	0.051	-0.012	0.991
manufacturing	-0.0196***	0.006	-3.096	0.002
construction	-0.0313**	0.013	-2.370	0.018
services1	0.0085	0.011	0.765	0.445

Table 8: The regression of $\hat{\varepsilon}_t^2$ - full results.

variable	**coef**	**std err**	**t**	**p-value**
const	0.0347***	0.003	10.557	0.000
C_fixed_term	0.0090***	0.001	13.058	0.000
log_income	-0.0019***	0.000	-5.591	0.000
male	-0.0041***	0.000	-17.433	0.000
tenure	-0.0002***	1.91e-05	-12.155	0.000
d1999	0.0005*	0.000	1.747	0.081
d2000	0.0005*	0.000	1.644	0.100
d2001	-0.0003	0.000	-0.996	0.319
d2002	-0.0007**	0.000	-2.257	0.024
white_colar	0.0013***	0.000	5.330	0.000
cadre	0.0025***	0.001	4.130	0.000
manager	-0.0047***	0.001	-5.317	0.000
North-East	-0.0013***	0.000	-5.223	0.000
Center	0.0002	0.000	0.916	0.360
South	0.0018***	0.000	6.069	0.000
Islands	-0.0003	0.000	-0.794	0.427
young_firm	0.0017***	0.000	7.015	0.000
agriculture	-0.0066***	0.002	-3.982	0.000
heavy_industry	-0.0017	0.001	-2.505	0.012
manufacturing	-0.0002	0.000	-0.837	0.402
construction	0.0023***	0.000	6.066	0.000
services1	0.0011***	0.000	3.460	0.001

Table 9: The regression of $\hat{\varepsilon}_{t-1}\hat{\varepsilon}_t$ - full results.

variable	**coef**	**std err**	**t**	**p-value**
const	0.0031	0.002	1.383	0.167
C_fixed_term	-0.0012	0.001	-1.568	0.117
log_income	-0.0006***	0.000	-2.742	0.006
male	0.0018***	0.000	11.332	0.000
tenure	-2.907e-05**	1.3e-05	-2.243	0.025
white_colar	0.0008***	0.000	5.051	0.000
cadre	0.0007*	0.000	1.746	0.081
manager	0.0031***	0.001	5.273	0.000
North-East	0.0003**	0.000	1.999	0.046
Center	-0.0004**	0.000	-1.987	0.047
South	-0.0010***	0.000	-5.057	0.000
Islands	-0.0003	0.000	-0.895	0.371
young_firm	-0.0002	0.000	-1.172	0.241