Optimal Taxation with Permanent Employment Contracts

Paweł Doligalski*

June 12, 2016

[most recent version]

Abstract

New Dynamic Public Finance describes the optimal income tax in the economy without private insurance opportunities. 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 the permanent employment. Such arrangement discourages the tax avoidance of the productive workers and hence allows the government to tax them more. 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.

*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, Warsaw Economic Seminar, SAEe and Econometric Society European Winter Meeting. 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.

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. 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 with a high firing cost discourage employers from laying off their employees, thus allowing firms to act as insurers. The government optimally outsources most of the insurance to the better informed firms and, depending on social objectives, can focus on redistribution. As a result, the optimal tax system is simple: in the model calibrated to Italy any reasonable constrained efficient allocation can be implemented with a tax schedule that depends exclusively on current consumption expenditure.[4] Such tax was proposed by various public finance economists in US.[5] It contrasts with the standard implementation of NDPF which involves time-varying taxation of labor income and capital income that depends on the whole history of past earnings.

The insurance within firms comes at a price. Permanent contracts reduce the random variation of income over the life-cycle, but they also allow firms to shift workers' compensation across time in order to minimize the workers' tax burden. 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.

There is strong empirical evidence of income shifting within firm, both for insurance and tax avoidance reasons. Guiso, Pistaferri, and Schivardi (2005) document that Italian firms insure workers by reducing variability of their income. Lagakos and Ordonez (2011) conduct a similar study for US and find that high-skilled workers obtain more insurance than low-skilled ones.[6] Kreiner, Leth-Petersen, and Skov (2015) describe the shifting of salaries within firms in response to the announced decrease of the top income tax rate in Denmark. Individuals affected by the reform shifted on average 10% of their labor income, although the effect is concentrated in a relatively small group of taxpayers that shift most of their salaries. In a companion paper, Kreiner, Leth-Petersen, and Skov (2014) focus on top management. Managers are most likely to shift income by retiming bonus payments, but delaying the regular wage income is also evident.

In my 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. At first I consider the frictionless labor market, where workers and firms can credibly promise not to terminate the employment relationship. I show that the full commitment between workers and firms severely restricts the redistributive power of the state. For instance, when workers are risk neutral, only progressive tax schedules are incentive-compatible. If the tax was

[3]Financial markets are also unlikely to observe individual productivity of every worker.

[4]By a reasonable allocation I mean the allocation that does not involve redistribution of income from the poor to the rich.

[5]The progressive consumption tax was advocated by Hall and Rabushka (1995) and Bradford (2000).

[6]Although there is no mandatory firing cost in US, Bishow and Parsons (2004) shows that between 1980 and 2000 on average 30% of employees in private establishments were covered by a voluntary severance pay provided by the employer. White collar workers are more likely to be covered, which can explain the gap in insurance between skill groups.

locally regressive, workers and firms would agree to randomize wages in order to reduce the average tax rate faced by the worker. Although the full commitment case is not realistic, it clearly shows that a reduction in contracting frictions between workers and firms exacerbates the tax avoidance and restricts redistribution. This observation will be useful in understanding the optimality of fixed-term contracts in the case without commitment on the labor market. The characterization of the full commitment case also sheds light on the generality of the optimal tax rate formulas expressed with sufficient statistics. Chetty and Saez (2010) show that the sufficient statistics formula for the linear income tax is valid also when workers have access to private insurance, as long as this insurance does not suffer from moral hazard. My results indicate that their finding cannot be generalized to the non-linear income tax. The optimal tax formulas derived by Diamond (1998) and Saez (2001) typically yield the U-shaped tax schedule, with marginal tax rates decreasing below the mode income. If the risk aversion of workers is sufficiently low, they could exploit the tax regressivity on low income levels by wage randomization.[7]

The main part of the paper is devoted to the frictional labor market, where neither workers nor firms are unable to commit to maintain the employment relationship in the future. Workers are free 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 heterogeneous only in labor productivity. In the opposite case, 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 effectively lock workers in the company. As workers no longer have incentives to quit, firms can offer them full consumption insurance.

I show that the workers that pay the highest taxes should always have permanent contracts and enjoy full consumption insurance. If they had fixed-term contracts instead, assigning them permanent contract would lead to a Pareto improvement for any tax system in place. The intuition is simple: with permanent contract, paying high taxes becomes more attractive. If this reform induced some other workers to change their behavior, they would end up contributing more resources to the governments budget. It could, nevertheless, be suboptimal to equip all workers with permanent contracts. When the government cares most about the initially least productive, these workers

[7] Another striking example of the difference in optimal tax system with and without private insurance is the top tax rate. Consider the economy with a bounded productivity distribution. In the standard Mirrlees (1971) model the optimal top tax rate is non-positive. In contrast, in the model with full commitment on the labor market the optimal tax rate is positive when the government wants to redistribute towards the less productive 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.[8] I complement this literature by showing how fixed term contracts influence individual responses to income taxation.

How to implement the optimal allocation with taxes? When all workers optimally have permanent contracts and full consumption insurance, they should face only the redistributive tax based on consumption expenditures.[9] 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. I show that the tax is governed by a well-understood Saez (2001) formula from the static Mirrlees (1971) model.[10] 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. Hence, it would be dealt with by firms. When tax payments increase progressively with consumption expenditures, the tax schedule can depend only on current consumption expenditures - no history dependence is required. Furthermore, 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. 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, all allocations at the Pareto frontier which do not

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

[9] The tax system described in this paragraph implements the optimum, unless the planner wants to redistribute income from the bottom to the top. In such unusual cases this implementation can yield a suboptimal outcome.

[10] Recall that the optimal tax system with full commitment on the labor market is not consistent with the Saez (2001) formula due to the threat of wage randomization. Introduction of the limited commitment on the side of workers is enough to prevent the wage randomization and recover the sufficient statistics formula.

involve redistribution from the bottom to the top 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%.[11] 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 Ábrahám, Koehne, and Pavoni (2016) show that hidden asset trades reduce the optimal progressivity of labor income tax. I find that income shifting, which is related to asset trades, reduces the redistribution possible via the income tax.

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 (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 continu-

[11] 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 4). The permanent contracts regime improves NDPF relative to the laissez-faire by more than 50%.

ous 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 to 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 optimum with full commitment on the labor market is characterized in Section 3. Section 4 characterizes the constrained efficient allocation with limited commitment. Implementation with a 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

In this section I describe the structure of the labor market, define the equilibrium and set up the optimal taxation problem.

2.1 Workers and firms

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 $\beta = R^{-1}$ is a discount factor.

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.

The labor market operates in the following way. Workers enter the market after their initial productivity is drawn. Firms make them offers which specify the labor supply and the labor income at each productivity history. I assume no search friction - all workers see all the offers immediately - which leads to a Bertrand competition between firms for workers. Once the contract is signed, the terms of the contract cannot be changed.[12] However, the contract can be terminated at will by both parties. At any point in time workers are free to leave their current employer and start a new job elsewhere. Workers face no mobility cost. On the other hand, firms can fire their employees in any period subject to the specified firing cost.[13] 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* ($f = 0$).

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. The length of history h - the number of productivity draws it contains - is denoted by $|h|$. 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 of length t that happen with positive probability. 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 denoted by $\mu(s \mid h)$. I assume that each initial type faces the productivity risk: $\forall_{\theta \in \Theta_1} \forall_{h \in \mathcal{H}_{\bar{t}}(\theta)}\ \mu(h \mid \theta) < 1$.

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.

Now we can specify the payoffs of agents. The expected utility of a worker at the history $h \in \mathcal{H}$, given the allocation (c, y, n) is

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

[12] It is an important assumption. If firms were unable to commit to the terms of the contract, the equilibrium would involve no private insurance regardless of the firing cost.

[13] 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.

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

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

I denote the expected payoffs of workers from the *ex ante* perspective by dropping the superscript: $\mathbb{E}U(c, n) \equiv \sum_{\theta \in \Theta_1} \mu(\theta) \mathbb{E}U_\theta(c, n)$, and analogously for firms.

2.3 The social planner

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 and individual output θn. The distinction between the observable labor income y and the unobservable output θn is realistic and crucial for modeling the insurance and the tax avoidance within firm. If the worker was paid his output in every period, there would be no insurance on the labor market. If the planner observed not only labor income, but also output, firms would not be able to use income shifting to reduce the tax burden of workers. The social planner sets up a mechanism which governs the allocation of resources in the economy. By the revelation principle, without the loss of generality we can focus our attention 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 in each period collects type reports of workers and assigns them consumption levels, labor incomes and firing costs. The agents' reports and the unobserved labor supply are determined in the equilibrium corresponding to the chosen mechanism. From now on I fix the message space at $\mathcal{H}$ and identify a given mechanism with its outcome functions (c, y, f).

Let's formalize the possible reporting behavior of agents. 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 distribution assigning all the probability mass to the truthful reporting strategy r^* is denoted by σ^*.

The expected utility of a worker at the history h, given outcome functions (c, y), a pure reporting strategy r and a corresponding labor allocation n_r is $\mathbb{E}U_h(c \circ r, n_r)$, 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 $\mathbb{E}\pi_h(y \circ r, n_r)$. Therefore, the reporting strategy directly affects the outcomes that are assigned by the mechanism. The payoffs of a worker and a firm from the mixed reporting strategy $\sigma \in \Delta_{\mathcal{R}}$ and the corresponding labor allocation $n_\sigma = \{n_r : \mathcal{H} \to \mathbb{R}_+\}_{r \in \mathcal{R}}$ at history h are $\sum_{r \in \mathcal{R}} \sigma(r) \mathbb{E}U_h(c \circ r, n_r)$ and $\sum_{r \in \mathcal{R}} \sigma(r) \mathbb{E}\pi_h(y \circ r, n_r)$. Note that in the case of the mixed reporting strategy, the labor supply allocation is allowed to vary with the selected pure reporting strategy.

2.4 Equilibrium

Since all firms are identical, there is no gain from workers changing employers. Hence, without the loss of generality, I focus on the equilibria without separations. The following lemma describes the conditions such equilibria have to satisfy.

Lemma 1. *(σ, n_σ) is such that neither a worker nor a firm has incentives to terminate the employment relationship if and only if*

$$\forall_{r\in\mathcal{R}\ s.t.\ \sigma(r)>0}\forall_{h\in\mathcal{H}} - f\left(r\left(h\right)\right) \leq \mathbb{E}\pi_h\left(y\circ r, n_r\right) \leq 0. \tag{3}$$

The worker has incentives to leave if the employer makes positive profits on him. If that happens, a competing firm could offer the worker a better deal, while still being profitable. On the other hand, the firm has incentives to fire the worker if the expected loses are greater than the firing cost. The limited commitment constraints (3) prevent both deviations.

Definition 3. (σ, n_σ) is the equilibrium of mechanism (c, y, f) if *(i)* (σ, n_σ) satisfies (3) and *(ii)* there is no other $(\sigma', n'_{\sigma'})$ which satisfies (3) and additionally $\sum_{r\in\mathcal{R}}\sigma'(r)\mathbb{E}U(c'\circ r, n'_r) > \sum_{r\in\mathcal{R}}\sigma(r)\mathbb{E}U(c\circ r, n_r)$ and $\sum_{r\in\mathcal{R}}\sigma'(r)\mathbb{E}\pi(c'\circ r, n'_r) \geq \sum_{r\in\mathcal{R}}\sigma(r)\mathbb{E}\pi(c\circ r, n_r)$.

The equilibrium reporting strategy and the labor supply allocation are determined by the payoff maximizing behavior of workers and firms, given the competition and the limited commitment on the labor market. Specifically, there can be no other $(\sigma', n'_{\sigma'})$ which is consistent with the limited commitment constraints, yields not lower profits to firms and strictly greater utility to workers. The following lemma describes the set of equilibria of a mechanism.

Lemma 2. *The set of equilibria of mechanism (c, y, f) is*

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

where maximization in subject to the limited commitment constraints (3) and the zero profit condition

$$\forall_{\theta\in\Theta_1}\sum_{r\in\mathcal{R}}\sigma(r)\mathbb{E}\pi_\theta\left(y\circ r, n_r\right) = 0. \tag{4}$$

Since workers observe all offers, the competition between firms for workers drives profits to zero. Notice that 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.

Definition 4. The mechanism (c, y, f) *implements* allocation (c, y, n) if $(\sigma^*, n) \in \mathcal{E}(c, y, f)$, i.e. if labor supply allocation n and the truthful reporting strategy constitute the equilibrium of the mechanism (c, y, f).

Note that the above notion of implementation does not require (σ^*, n) to be the unique equilibrium of the mechanism. The optimal mechanisms generally have multiple equilibria, some of which involve untruthful reporting. I implicitly assume that if there exists a truthful equilibrium, the agents will choose it.[14]

2.5 The planner's problem

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

$$\max_{\substack{c:\mathcal{H}\to\mathbb{R}_+\\ y:\mathcal{H}\to\mathbb{R}\\ f:\mathcal{H}\to\{0,\bar{f}\}}} \sum_{\theta\in\Theta_1} \lambda(\theta)\,\mu(\theta)\,\mathbb{E}U(c,n)\,, \tag{5}$$

where λ is the non-negative Pareto weight with the expected value of 1: $\sum_{\theta\in\Theta_1}\lambda(\theta)\,\mu(\theta)=1$. The optimization is subject to the resource constraint

$$\sum_{h\in\mathcal{H}} R^{1-|h|}\mu(h)\,(y(h)-c(h))\geq 0 \tag{6}$$

and the equilibrium constraint

$$(\sigma^*,n)\in\mathcal{E}(c,y,f)\,. \tag{7}$$

The equilibrium constraint means that the chosen mechanism (c,y,f) implements the allocation (c,y,n). It incorporates the usual incentive compatibility constraints that prevent type misreporting. Note that the untruthful equilibria in $\mathcal{E}(c,y,f)$ correspond to the binding incentive constraints.

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 maintain the employment relationship. However, I do not allow firms and workers to contract before the initial productivity draw. If contracting behind the veil of ignorance was allowed, as in Golosov and Tsyvinski (2007), firms could redistribute between initial types. Here redistribution within firm is prevented, as any labor contract involving cross-subsidization allows competitors to profitably steal the worker that is paid less than his product.

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

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

Since firms and workers can credibly commit to maintain the employment relationship, we can drop the limited commitment constraints. This means that the firing cost does not influence the

[14] 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 $\mathcal{E}(c,y,f)$.

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. Suppose that at the equilibrium reporting strategy the expected lifetime income of initial type h_1 is $Y(h_1)$. The necessary and sufficient condition for the optimal labor supply of each initial type is to equalize the marginal cost of output across all histories and reporting strategies

$$\forall_{r\in\mathcal{R}}\forall_{h\in\mathcal{H}}\frac{v'\left(n_r\left(h\right)\right)}{\theta\left(h\right)}\equiv\phi_{h_1}(Y(h_1)). \tag{8}$$

Under full commitment the equilibrium allocation of labor supply produces the expected lifetime income Y at the minimal disutility cost. The output and the labor income agree in expectations over the lifetime of a worker, which is captured by the zero profit condition, but do not have to coincide at every history. It means that the firm can shift worker's income across time and productivity histories. The output produced by worker of initial type θ at some history $h\in\mathcal{H}(\theta)$ can be paid to him at any other history $h'\in\mathcal{H}(\theta)$, as long as the zero profit condition (4) holds. Such an unrestricted income shifting is possible only because of the full commitment of firms and workers.

Let's denote by $n_Y^{FC}:\mathcal{H}\to\mathbb{R}_+$ the labor supply allocation which satisfies the optimality condition (8) and generates the expected lifetime income Y. We can use it to construct the indirect utility function that captures the expected utility of some initial type θ from lifetime consumption C and lifetime labor income Y:

$$V_\theta\left(C,Y\right)\equiv\bar{\beta}u\left(C/\bar{\beta}\right)-\sum_{h\in\mathcal{H}(\theta)}\beta^{|h|-1}\mu(h\mid\theta)v\left(n_Y^{FC}(h)\right)$$

where $\bar{\beta}\equiv\sum_{t=1}^{\bar{t}}\beta^{t-1}$. Notice that $V_\theta(C,Y)$ implicitly assumes that the worker enjoys full consumption insurance, while the labor supply is chosen in order to minimize the disutility cost of producing the lifetime income Y. By the theorem below, we can use this indirect utility function to simplify the taxation problem under full commitment.

Theorem 1. *Under full commitment on the labor market, all workers enjoy full consumption insurance and the planner's problem can be expressed as*

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

subject to the resource constraint

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

and the incentive-compatibility constraints

$$\forall_{\theta,\theta'\in\Theta_1}\bar{\beta}u(C(\theta)/\bar{\beta})-Y(\theta)\phi_\theta(Y(\theta))\geq\bar{\beta}u(C(\theta')/\bar{\beta})-Y(\theta')\phi_\theta(Y(\theta)). \tag{9}$$

Under full commitment on the labor market firms are ideal insurers of workers. Firms are driven by competition to provide workers with the maximal utility attainable without making losses. Moreover, they are better informed than the planner. The optimal mechanism makes use of firms to provide full consumption insurance to workers.

By Theorem 1, the planner chooses only the lifetime consumption and lifetime labor income of each initial type. Since all agents enjoy constant consumption, a lifetime consumption fully determines the consumption at each history. Furthermore, because of full commitment, the allocation of labor in equilibrium depends only on the expected lifetime income and not on labor income on any particular history. No matter how the labor income is structured by the planner, the firm will always allocate labor to histories in a way that minimizes the total disutility cost of lifetime production.

The incentive compatibility constraints (9) are not standard. The reason is that in the optimum the worker is never tempted by reporting some other type with certainty. In the proof of Theorem 1 I show that if that was the case, then the worker would be strictly better off with mixing between this reporting strategy and truth-telling. The mixed reporting strategy is better, because under full commitment the firm can equalize the marginal cost of production across pure reporting strategies over which the worker randomizes. Since the disutility from labor is strictly convex, the worker strictly gains from this labor smoothing across reporting strategies. Consequently, only the incentive constraints corresponding to the mixed strategies can bind in the optimum. Condition (9) means that the gain from a marginal increase in probability of reporting θ' when the true type is θ is non-positive. It is a necessary and sufficient condition for truth-telling when workers can use mixed reporting strategies.

Since the incentive constraints with respect to all pure reporting strategies need to be slack, the truth-telling places a tight constraint on implementable allocations. To see this, define a lifetime tax $T(Y(\theta)) \equiv Y(\theta) - C(\theta)$ and its marginal rate $T'(Y(\theta)) \equiv 1 - \frac{\phi_\theta(Y(\theta))}{u'(C(\theta)/\beta)}$.

Proposition 1. *Under full commitment on the labor market, the mechanism* $(C(\theta), Y(\theta))_{\theta \in \Theta_1}$ *implements truth-telling only if* $(1 - T'(Y))u'\left(\frac{Y - T(Y)}{\beta}\right)$ *is non-increasing in* Y*.*

Full commitment on the labor market restricts the regressivity of the tax schedule. For instance, when workers are risk neutral, only progressive (i.e. convex) tax schedules are implementable. Suppose on the contrary that the tax schedule is strictly regressive on some income interval $[Y, Y']$. It means that the marginal tax rate at Y is greater than the average tax rate at this interval:

$$T'(Y) > \frac{T(Y') - T(Y)}{Y' - Y}. \tag{10}$$

By substituting the definitions of the tax and the marginal tax rate one can show that the incentive compatibility constraint (9) is violated. Suppose that the worker with income Y marginally increases output. How should a firm compensate the worker? The additional income can be paid with certainty and taxed at the marginal tax rate. Alternatively, the firm can compensate the worker with additional income $Y' - Y$ which is paid with probability low enough such that the firm makes no losses. This additional income is taxed at the average tax rate at the income interval $[Y, Y']$. Whenever the average rate is lower than the marginal rate, i.e. whenever the tax schedule is strictly regressive, the risk-neutral worker strictly prefers random compensation. For risk averse workers such income randomization is naturally less attractive, hence some degree of regressivity is still possible.

Chetty and Saez (2010) show that the sufficient statistics formula for the optimal linear income tax is valid also in the presence of private insurance, as long as private insurers do not suffer from moral hazard. In my framework the private insurance is free of moral hazard, since firms observe workers' types. Hence, by Proposition 1 the result of Chetty and Saez (2010) does not generalize to a non-linear income tax. The sufficient statistic formula for the optimal non-linear income tax may prescribe a locally regressive tax. In fact, the optimal taxation literature typically recommends the U-shaped tax schedules, with tax rates decreasing below the mode income (see Diamond (1998); Saez (2001)).[15] If the labor market operated under full commitment and the risk aversion was sufficiently low, such tax would induce the tax avoidance via income randomization at low levels of income. As a result, the tax revenue would be lower than predicted.

Another way to see that the optimal tax rate under full commitment is qualitatively different than the optimal tax rate without private insurance is to consider the top tax rate. In the static Mirrlees (1971) model this rate is always non-positive. In the model with the full commitment on the labor market this rate will be positive when the incentive constraint of the top type binds. That is the case because decreasing labor supply of the top type reduces his utility from the marginal deviation to a mixed reporting strategy.

Proposition 2. *Suppose that workers live for one period ($\bar{t} = 1$) and the planner is utilitarian. In the optimum, the labor supply of the top type is distorted downwards.*

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, subject to the firing cost. From the previous section we know that under full commitment on the labor market the set of implementable allocations is severely constrained by the possibility of using mixed reporting strategies. Without commitment on the side of workers mixed strategies are much less powerful and we can focus exclusively on pure reporting strategies.

Lemma 3. *Under limited commitment on the labor market, the payoff from any mixed reporting strategy is dominated by the payoff from some pure reporting strategy.*

With full commitment workers could smooth labor across the different pure reporting strategies over which they were mixing. At some pure strategies the firm made positive profits, at others - suffered losses. Without commitment on the workers' side such arrangement is not sustainable, as workers have incentives to leave the firm if it makes strictly positive profits. Hence, the limited commitment of workers prevents firms from reducing a tax burden via the wage randomization.

[15] Such recommendations are drawn from the static Mirrlees (1971) model. It is a special case of my framework, when workers live for only one period ($\bar{t} = 1$) and the commitment on the labor market is limited.

Without commitment on the labor market the type of labor contract matters, since the high firing cost prevents firms from dismissing their workers. The following two subsections describe the optimal allocation when the planner is restricted to use only fixed-term or only permanent contracts. Finally I describe the optimal choice of the contract type.

4.1 Only fixed-term contracts

Suppose that the planner assigns fixed-term contracts to workers at each history: $\forall_{h\in\mathcal{H}} f(h) = 0$.

Lemma 4. *Under fixed-term contracts, in any equilibrium (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 to maintain the employment. 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 4 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 (7) as

$$\forall_{r\in\mathcal{R}} \mathbb{E}U(c, \bar{n}(r^*)) \geq \mathbb{E}U(c \circ r, \bar{n}(r)), \text{ where } \forall_{h\in\mathcal{H}} \bar{n}(r(h)) = \frac{y(r(h))}{\theta(h)}. \tag{11}$$

This is exactly the incentive-compatibility constraint considered by NDPF. Since the firms do not insure their workers, the government steps in with the tax system which both redistributes and insures. 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 (2016) provide the detailed characterization of the optimal labor wedges.

4.2 Only permanent contracts

Suppose that the planner uses only permanent contracts: $\forall_{h\in\mathcal{H}} f(h) = \bar{f}$. 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 equilibria in the similar settings were characterized by Harris and Holmstrom (1982) and Krueger and Uhlig (2006).[16] The firm overcomes a worker's commitment problem by backloading labor compensation,

[16] 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.

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 full commitment on the labor market. The planner can implement it under permanent contracts.*

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.[17] 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 this result 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 or tenure dependent taxation.

Corollary 3. *For any Pareto weights $\{\lambda(\theta)\}_{\theta\in\Theta_1}$, the optimum without commitment on the labor market yields weakly higher social welfare than the optimum with full commitment. The relation is strict if the optimum with full commitment features binding incentive constraints.*

The first statement is a simple implication of Theorem 2. The second statement comes from the fact that the full commitment optimum the binding incentive constraints corresponding to the mixed strategies. However, Lemma 3 shows that without commitment on the labor market these constraints become slack. Hence, without commitment between firms and workers the redistributive planner is less constrained by tax avoidance and can achieve higher social welfare.

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, effectively 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.[18] For an example of such a situation, see Lemma A.2 in the Appendix.

[17] 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.

[18] A dual labor market allocation is preferable, because fixed-term contract prevents income shifting of the deviating type. However, in some cases the limited commitment constraints of worker are enough to prevent the income shifting. 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.2.

4.3 Who should have permanent contract?

In the following two subsections I investigate which workers should have permanent contracts, and which fixed-term contracts. Theorem 3 states that workers that pay the highest taxes should have permanent contracts.

Definition 5. *An initial top taxpayer* is a type that belong to

$$\arg\max_{\theta\in\Theta_1} \sum_{s\in\mathcal{H}(\theta)} R^{1-|s|}\mu(s\mid\theta)\,(y(s)-c(s)).$$

Theorem 3. *Initial top taxpayers optimally have permanent contracts and full consumption insurance.*

Assigning permanent contracts allows the planner to provide more insurance and save resources, but it also increases the incentives of other workers to misreport. However, there are some types that can be mimicked without a loss for the planner: initial top taxpayers. If any other initial worker decides to report that he is a top taxpayer, he will end up contributing more resources to the planner's budget. Hence, simply assigning permanent contracts to top taxpayers is a Pareto improving reform. 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. 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 4. *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. If a planner cares only about insurance, it is optimal to use only permanent contracts.

4.4 Who should have fixed-term contract?

Take some allocation (c,y,n) with corresponding contract assignment f where the initial type $\underline{\theta}$ has permanent contract. Consider an alternative contract assignment f' where the worker at the history $\underline{\theta}$ (and all histories that follow) receives a fixed-term contract and the contract types of other workers are unchanged. Denote the best allocation of consumption and labor, conditional on contract assignment f', by (c',n'). Let's write the social welfare function as $W(c,n)\equiv\sum_{\theta\in\Theta_1}\lambda(\theta)\mu(\theta)\mathbb{E}U_\theta(c,n)$. We can decompose the welfare impact of switching contract type into three components, capturing the change in efficiency, redistribution and insurance:

$$W(c',n')-W(c,n) =$$
$$\underbrace{W(c',n')-W(c_2,n)}_{\Delta^{efficiency}} + \underbrace{W(c_2,n)-W(c_1,n)}_{\Delta^{redistribution}} + \underbrace{W(c_1,n)-W(c,n)}_{\Delta^{insurance}}.$$

Consumption allocation c_1 involves a consumption risk of the fixed-term contract of $\underline{\theta}$, but keeps the present value of consumption of each initial type at the same level as the original allocation c. Hence, $\Delta^{insurance}$ captures the welfare loss due to missing insurance within firm. Consumption allocation c_2 contains both the consumption risk and the change in transfers between initial types. The transition to fixed-term contract relaxes incentive-compatibility constraints between initial types for two reasons: the consumption of $\underline{\theta}$ is more volatile and income shifting is no longer possible. Thus, $\Delta^{redistribution}$ captures the welfare impact of the change in redistribution. Finally, (c', n') correspond to solving the standard government's problem given the contract allocation f'. Note that only at this stage the adjustment to labor supply are allowed. Hence, $\Delta^{efficiency}$ expresses the welfare gain from optimal adjustment of both consumption and labor along the incentive-compatibility constraints. For the details of this decomposition, see Definition 8 in the Appendix.[19]

Lemma 5. *$\Delta^{insurance} \leq 0, \Delta^{efficiency} \geq 0$. $\Delta^{redistribution} \geq 0$ if $\underline{\theta} \in \arg\max_{\theta \in \Theta_1} \lambda(\theta) u'(c_1(\theta))$ and $\Delta^{redistribution} \leq 0$ if $\underline{\theta} \in \arg\min_{\theta \in \Theta_1} \lambda(\theta) u'(c_1(\theta))$.*

Lemma 5 determines the signs of the decomposition terms: the switch from permanent to fixed-term contract leads to a utility loss due to lower insurance and utility gain due to labor adjustment. The sign of the redistribution component depends on the desired direction of redistribution. Fixed-term contract improves redistribution if assigned to a recipient of government transfers rather than a net taxpayer.

Lemma 5 suggests that there are two channels that may lead to the optimality of fixed-term contracts: redistribution and efficiency. I explore these channels in the two propositions below.

Assumption 1. *The distribution of productivity has full support: $\forall_{h \in \Theta^{\bar{t}}} \mu(h) > 0$, and satisfies the first-order Markov property: for any $h \in \mathcal{H}$ such that $|h| > 2$ and any $s \in \mathcal{H}_{|h|-2}$ it is true that $\mu(h) = \mu(s, h_{|h|-1}, h)$.*

Proposition 3. *Suppose that (i) Assumption 1 holds and additionally the distribution of productivities is independent of the initial productivity draw, (ii) workers are risk neutral: $U(c,n) = c - v(n)$, (iii) $\underline{\theta}$ is the initial type with the lowest productivity and has a positive lifetime income: $\max_{s \in \mathcal{H}(\underline{\theta})} y(s) > 0$, (iv) the planner is Rawlsian: $\forall_{\theta \neq \underline{\theta}} \lambda(\theta) = 0$. Assigning fixed-term contract to type $\underline{\theta}$ is welfare improving.*

To understand this proposition, consider a simple example with two initial types $\left(\Theta_1 = \left\{\underline{\theta}, \overline{\theta}\right\}, \overline{\theta} > \underline{\theta}\right)$. The planner wants to maximize the utility of the low type and hence will redistribute from $\overline{\theta}$ to $\underline{\theta}$. What limits redistribution is ability of $\overline{\theta}$ to mimic the other type. 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 instead, 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 in the future, allowing the mimicking worker to reduce the future labor supply. This income shifting implies that the high type gains more from misreporting when the low type has permanent contract rather than fixed-term contract.

[19] Doligalski and Rojas (2016) use the similar decomposition of welfare change into redistribution and efficiency components in the static Mirrlees (1971) framework with an informal sector.

The simplifying assumption of risk neutrality means that the planner cares only about the redistribution and not about the insurance ($\Delta^{insurance} = 0$). Since there is no utility loss from volatile consumption, the incentive constraints that prevent redistribution are relaxed only because of the prevented income shifting. With two initial types $\left(\Theta_1 = \left\{\underline{\theta}, \overline{\theta}\right\}, \overline{\theta} > \underline{\theta}\right)$, we can express the redistribution gain explicitly as

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

where $\tilde{n}$ is labor allocation of the mimicking high type when $\underline{\theta}$ has permanent contract, while $\hat{n}$ is the labor allocation when $\underline{\theta}$ has fixed-term contract. In line with the intuition above, the disutility from labor of the mimicking type is higher if the other type has fixed-term contract. Thus, $\Delta^{redistribution}$ is strictly positive. Since there is no utility loss due to missing insurance and $\Delta^{efficiency}$ is always non-negative, the overall welfare impact of switching contract of $\underline{\theta}$ is positive. Although the risk neutrality is a strong assumption, we can expect this result to hold also for moderate risk aversion, when $\Delta^{insurance}$ is sufficiently small.

Proposition 4. *Suppose that (i) Assumption 1 holds, (ii) there is some initial type $\underline{\theta}$ with permanent contract that supplies no labor: $\forall_{h \in \mathcal{H}(\underline{\theta})} n(h) = 0$ and faces downward labor distortions in future periods: $\exists_{h \in \mathcal{H}(\underline{\theta}) \setminus \underline{\theta}}$ s.t. $\theta(h) u'(c(h)) > v'(0)$. Assigning fixed-term contract to type $\underline{\theta}$ is welfare improving.*

Proposition 4 shows that fixed-term contracts can improve the allocation of labor. Suppose that the distortions under permanent contracts are so severe that the some type has no lifetime earnings.[20] Notice that $\Delta^{insurance}$ is zero also in this case, although this time we do not impose risk neutrality. Since $\underline{\theta}$ does not supply labor, there is no need for volatile consumption. Moreover, $\Delta^{redistribution}$ is zero as well, for there is no scope for income shifting. The low type can gain only though the efficiency considerations.

Under permanent contracts, the planner discourages misreporting by reducing labor income of $\underline{\theta}$ at all histories. This is the case, because the output produced initially can be paid to the worker at any future history. Such income shifting is not possible with fixed-term contract. Under fixed-term contract the planner can lift some of the future distortions and generate additional resources, achieving $\Delta^{efficiency} > 0$. For instance, in the simplest iid case the classical 'no distortion at the top' result extended to the dynamic setting says that it is suboptimal to distort labor supply of the most productive type after any history. The planner should lift distortions of the most productive fixed-term workers. Note that not all distortions should be lifted, since they serve insurance purpose.

Hopenhayn and Rogerson (1993) claim that high severance payments cause labor misallocation. Their finding relies on the contractual friction: workers are paid a market wage in every period. Lazear (1990) shows that if instead labor contracts were flexible, the firm could design a compensation structure that nullifies the adverse effect of the firing cost on employment. In this paper the logic of Lazear (1990) holds: the high firing cost of permanent contract does not discourage

[20] Optimum under permanent contracts has this feature when $v'(0) > 0$, $\mu(\underline{\theta})$ is sufficiently low and the planner wants to redistribute to $\underline{\theta}$.

firms from hiring workers.[21] The firing cost does, on the other hand, encourage firms to offer a compensation structure that minimizes workers' tax burden. The government can prevent firms from doing so either by introducing additional tax distortions or by promoting fixed-term contracts. Proposition 4 identifies the case when the latter is preferable.

5 Simple fiscal implementation

Dynamic optimal taxation literature suffers from very complicated tax implementations. I tackle this problem in my framework by considering a restricted taxation problem. The restricted problem is attractive for a few reasons. Its solution can be described with the well understood Saez (2001) formula from the static Mirrlees model. This solution can implemented with a simple tax system, which, in the most favorable case, depends exclusively on current consumption expenditures. Furthermore, the restricted problem provides a tight lower bound on attainable welfare. In fact, in the next section I show numerically that for the typical social welfare functions the solution to the restricted problem coincides with the unrestricted optimum.

This section is structured as follows. First, I formalize the notion of the tax system and fiscal implementation. Then I consider the optimum with full consumption insurance: I define the restricted taxation problem and show that its solution can be implemented with simple tax system. Finally, I discuss the implementation of the dual labor market allocation.

5.1 The tax system

In the previous sections I characterized the optimal direct mechanism. This section is concerned with an indirect mechanism - a tax system. The tax can depend on all observables: history of labor income y, asset trades a and type of contract f as well as age t.

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

We can define the set of equilibria corresponding to the tax system. Firms and workers, who take the tax system as given, optimize with respect to labor supply, savings, the type of labor contract as well as compensation structure. The tax system affects the equilibrium by modifying the budget constraint of workers.

Lemma 6. *The set of equilibria given the tax system T is*

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

[21] In my framework there is no gain from workers switching jobs, as all firms are identical. However, even if there were efficient separations, high firing cost may still be efficient. Postel-Vinay and Turon (2014) show that firms facing high firing cost can persuade their workers to leave with a generous severance packages. Thanks to the high firing cost, the firm internalizes the worker's utility loss from separation.

subject to the zero profit condition and the limited commitment constraints

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

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

the sequence of budget constraints

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

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

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 $(c,y,n,a,f)\in\tilde{\mathcal{E}}(T)$.

5.2 The case of full consumption insurance

Suppose that it is optimal to assign permanent contracts and full consumption insurance to all workers. From Theorem 2 we know that the optimum under full commitment on the labor market provides a lower bound on welfare that can be achieved in the no commitment case. Furthermore, by Lemma 3 we know that without commitment the workers can no longer gain by deviating from truth-telling with mixed reporting strategies. Hence, I construct the restricted taxation problem by considering a full commitment problem, as in Theorem 1, with the incentive compatibility constraints that need to be satisfied only with respect to pure reporting strategies.

Definition 7. A *restricted taxation problem* is

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

subject to the resource constraint

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

and the incentive-compatibility constraints in pure strategies

$$\forall_{\theta,\theta'\in\Theta_1}V_\theta(C(\theta),Y(\theta))\geq V_\theta(C(\theta'),Y(\theta')).\tag{12}$$

Lemma 7. *A solution to the restricted taxation problem is implementable under permanent contracts.*

Lemma 7 means that the restricted taxation is a relevant lower bound for welfare in the unrestricted case, as there exists a direct mechanism that implements it. Note that the incentive constraints are tighter in the restricted problem than in the unrestricted problem. In the restricted problem the indirect utility function $V_\theta(C,Y)$ implicitly incorporates the optimality condition (8), which means that the marginal cost of output are equalized at every history. In the unrestricted case with permanent contracts the marginal cost of output needs to be only non-increasing over time. If it was increasing, the firm could shift the workers labor backwards in time, thereby reducing the overall disutility cost of labor. However, the marginal cost of output may be strictly decreasing over time, since the limited commitment of workers prevents the firm from shifting the labor forward. Recall that in this subsection we assume that the full consumption insurance for all workers is optimal. Then the solution to the restricted problem fails to reach optimum only if the optimal allocation features the decreasing marginal cost of output along some history. Such a distortion of labor supply may be optimal only if it relaxes the incentive compatibility constraints, which in turn happens only if the mimicking type prefers to produce later rather than earlier.

Suppose that the productivity process exhibits the mean reversion and the planner wants to redistribute from the initial low productivity type to the initial high productivity type. Then, by lifting the labor supply of the initial high type, the planner discourages the initial low type from misreporting. On the other hand, the planner cannot gain from a similar distortion while redistributing from the high to the low type, as lifting the initial labor supply of the low type would only encourage the misreporting of the high type. According to this intuition, we can expect the solution to the restricted problem to reach the unrestricted optimum when the planner wants to redistribute towards the less productive workers. This conjecture is true in the calibrated model considered in the next section.

The restricted problem is essentially the static Mirrlees (1971) model: the planner chooses lifetime consumption and lifetime income of each initial type subject to the initial incentive compatibility constraints in pure strategies. Similarly as in Section 3, let's denote by $T(Y)$ the net present value of taxes paid by an individual with a lifetime income Y: $T\left(Y(\theta)\right) \equiv Y(\theta) - C(\theta)$. Under the additional assumptions, we can express the solution to the restricted taxation problem with the modified Saez (2001) formula.

Assumption 2. *Define* $\hat{\mu}(\theta' \mid h_1) \equiv \sum_{h\in\mathcal{H}(h_1)} R^{1-|h|}\mu(h \mid h_1)/\bar{\beta}\mathbb{I}_{\theta(h)=\theta'}$*, where* $\mathbb{I}$ *is the indicator function. Take two initial types* $h_1, s_1 \in \Theta_1$*. If* $s_1 > h_1$*, then* $\hat{\mu}(\theta' \mid s_1)$ *first-order stochastically dominates* $\hat{\mu}(\theta' \mid h_1)$*.*

Assumption 3. Θ_1 *is an interval of real, non-negative numbers. The probability density function over* Θ_1 *is* $f(\theta)$ *and the cumulative distribution function is* $F(\theta)$*.*

Proposition 5. *Under Assumptions 2 and 3, if the implied lifetime income schedule* $Y(\theta)$ *is non-decreasing, the solution to the restricted taxation problem of an initial type* $\theta \in \Theta_1$ *satisfies*

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

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,* $\bar{\xi}(\theta) = \bar{\beta}^{-1}\sum_{h\in\mathcal{H}(\theta_1)} R^{1-|h|}\mu(h \mid \theta_1)\xi(h)$ *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,*

$\omega(\theta) = \lambda(\theta) u' \left(C(\theta)/\bar{\beta}\right)/\eta$ is the marginal social welfare weight of the initial type θ and η is the multiplier of the resource constraint.

Assumption 2 states that the distribution of productivity of initially higher types first-order stochastically dominates the distribution of productivity of the initially lower types. This assumption guarantees that the indirect utility function $V_\theta(C, Y)$ satisfies the Spence-Mirrlees single crossing property (see Lemma A.4 in the Appendix). The single crossing property intuitively means that the initial higher type is more eager to work on average over the whole lifetime than the initial lower type. If this property holds, any incentive-compatible lifetime income schedule is non-decreasing in the initial type. In order to apply the existing optimal tax formulas, we need to slightly modify the environment - Assumption 3 makes the initial distribution of types continuous. Given these assumptions, if the resulting income schedule is non-decreasing, we can express the optimal marginal tax rates with the formula derived by Saez (2001).

The 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.

Fiscal implementation of the allocation $\{C(\theta), Y(\theta)\}_{\theta\in\Theta_1}$ in the static Mirrlees (1971) model is simple. It is enough to have a tax system that depends on labor income according to the function $T_y \equiv T \circ Y^{-1}$. In the dynamic setting agents make multiple choices, hence the tax system needs to prevent multidimensional deviations. Furthermore, insurance within firms requires back-loading, which means that the labor income is not constant over the life-cycle. Implementing a full consumption insurance with a labor income tax would require a complicated, time-varying tax schedule. Instead, the tax can be based on consumption itself. Define *consumption expenditure* at the history h as the total income net of new savings: $x(h) \equiv y(h) + Ra(h^{-1}) - a(h)$. Consumption expenditure provides an attractive base for the redistributive tax since it is observable by the tax authority and stable over the workers' lifetime.

Take an allocation of lifetime consumption and labor income $\{C(\theta), Y(\theta)\}_{\theta\in\Theta_1}$, where the lifetime income is increasing in the initial type. Consumption expenditure of type θ is $x(\theta) \equiv Y(\theta)/\bar{\beta}$. Define a *consumption expenditure tax* as $\bar{T}_x \equiv \bar{T} \circ x^{-1}$, where $\bar{T}(\theta) \equiv (\bar{Y}(\theta) - C(\theta))/\bar{\beta}$ is the average tax paid by initial type θ. Extend this function to the non-negative real half-line with T_x, which is equal to $\bar{T}_x$ for values of x assigned for some type, and otherwise takes a prohibitively high value.

Theorem 4. *Take any allocation (c, y, n) and the corresponding allocation of lifetime consumption and income $\{C(\theta), Y(\theta)\}_{\theta\in\Theta_1}$ that is consistent with incentive compatibility constraints (12). Suppose that the borrowing limit is sufficiently high: $b \geq -\min_{h\in\mathcal{H}}\left\{\sum_{t=1}^{|h|} R^{1-|h|} y(h^t) - Y(h_1)\right\}$. If $\bar{T}_x$ is convex, then the allocation can be implemented with the tax system*

$$\forall_{t=\{1\ldots\bar{t}\}}\ T_t\left((y_k, a_k, f_k)_{k=1}^{t}\right) = T_x\left(x_t\right), \textit{where } x_t \equiv y_t + Ra_{t-1} - a_t.$$

If $\bar{T}_x$ is not convex, the allocation can be implemented with the tax system

$$\forall_{t=\{1\ldots\bar{t}\}}\ T_t\left(\left(y_k, a_k, f_k\right)_{k=1}^{t}\right) = T_x\left(x_t\right) + \alpha\left(x_t - x_1\right)^2,$$

where α is high enough such that $\bar{T}_x(x) + \alpha(x - x_1)^2$ is convex in x.

In the simplest case all we need for fiscal implementation is the time-invariant redistributive tax schedule based on current consumption expenditures.[22] Note that when the consumption expenditure tax is locally regressive (i.e. $\bar{T}_x$ is not convex), we need to add a corrective term that discourages variation in expenditures. Although the limited commitment prevents wage randomization, workers still can introduce a variation in expenditures over time. When a tax is regressive, such fluctuations would reduce their average tax burden and hence would attract workers with sufficiently low risk aversion. The corrective term convexifies the tax system and prevents this type of tax avoidance.[23] Conversely, when the consumption expenditure tax is progressive, no history dependence is required.

Corollary 5. *Consider Theorem 4. Suppose the borrowing is insufficient:*

$$b < -\min_{h\in\mathcal{H}}\left\{\sum_{t=1}^{|h|} R^{1-|h|} y(h^t) - Y(h_1)\right\}.$$

The allocation can be implemented with the tax system as in Theorem 4 combined with the governmental lending to workers.

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 taxes. Hence, the government in this setting can always relax the borrowing constraint of 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, in which the 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

[22] 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.

[23] Any schedule that is convex in x_t and equal to $T_x(x_1)$ for $x_t = x_1$ would work. For instance, workers in the period $t > 1$ can face a linear tax schedule or, in the simplest but perhaps the least realistic case, a tax that depends only on x_1 and not on current expenditures x_t.

persistence of productivity shocks. Although assigning permanent contracts to all workers is not always optimal, it is the feature of the utilitarian optimum in the calibrated model I consider in Section 7.

Existing tax codes bear a similarity to the consumption expenditure tax, in which savings are taxed only when they are spent on consumption. Unrealized capital gains are not taxed in US - income from the increased value of stocks is taxed only when the stocks are sold. Le Maire and Schjerning (2013) describe the Danish income tax for self-employed, which allows to retain earnings within the firm and pay them later in order to smooth the tax payments. Progressive consumption expenditure tax has been advocated by Bradford (2000). A flat tax proposed by Hall and Rabushka (1995) is a special case of such tax.[24]

5.3 The case of a dual labor market

When the dual labor market is optimal, the fiscal implementation is more complicated, as the tax system needs to insure the fixed-term workers against the productivity shocks. The tax system needs to separate the initial fixed-term workers and the initial permanent workers, hence the dependence of the tax system on the contract type may be required.[25] Nevertheless, the restricted taxation problem described in the previous subsection is still useful to describe the part of the tax system faced by the initial permanent workers. If none of fixed-term workers is tempted to mimic the initial permanent workers,[26] the restricted problem provides an implementable lower bound on welfare of the initial permanent workers. We only need to modify the resource constraint in order to capture the resource cost of transfers to the fixed-term workers.

The tax system of the fixed-term workers will follow findings of the NDPF literature. Specifically, it will involve savings taxation.

Proposition 6. *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.

[24] Bradford (2000) and Hall and Rabushka (1995) support the consumption tax because it encourages savings while allowing for redistribution. In my model savings play no productive role, since there is no capital. Instead, asset trades alleviate the contracting friction within the firm.

[25] Notice that fixed-term workers may be converted to permanent workers at some histories. The tax schedule they should then face after conversion is likely to be different than the tax schedule of the initial permanent workers. Hence, it is not enough that the tax depends only on a current contract type.

[26] As is likely to be the case, since fixed-term contracts facilitates redistribution toward the workers that have them.

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 provided by Bishow and Parsons (2004) that white collar workers are more frequently offered severance pay than blue collar workers.

6.1 Labor contracts in Italy

Italy in its modern history experienced a proliferation of distinct labor contracts.[27] 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.[28] Ichino, Polo, and Rettore (2003) provide 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.[29] 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

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

[28] As Ichino, Polo, and Rettore (2003) put it, "...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 (...), 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."

[29] 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.

time effects. Consider the following model:

$$\log(y_{ijt}) = \rho + W_{it}'\alpha + F_{jt}'\beta + M_{ijt}'\gamma + D_t'\delta + \epsilon_{ijt}, \tag{13}$$

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 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 (13)

$$\log\left(\frac{y_{ijt}}{y_{ijt-1}}\right) = \Delta W_{it}'\alpha + \Delta F_{jt}'\beta + \Delta M_{ijt}'\gamma + \Delta D_t'\delta + \Delta\epsilon_{ijt} \tag{14}$$

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 (14) 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.[30]

How does the variance of the error term in (13) 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

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

types have the same autocorrelation, then the risk ratio is simply given by the ratio of variances of errors from the differenced equation (14).

6.3 Data

The data comes from *Work Histories Italian Panel* (WHIP), a sample of administrative records of Italian employment histories.[31] 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 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 (14) is estimates with OLS separately for each type of contract.[32] 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

[31] Work Histories Italian Panel is a database of work histories developed thanks to the agreement between INPS and University of Turin. For more information, see http://www.laboratoriorevelli.it/whip.

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

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 (14). The main results of this regression are reported in Table 2, the full results and auxiliary estimates are reported in Appendix B.

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.

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

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.

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[33]. 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 figure 1). Under assumption of no borrowing in the data, the income process is informative about the productivity for all age/contract type groups.[34]

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. Figure 1 presents the estimation results. Income is reported in euros per year at the 2004 prices.

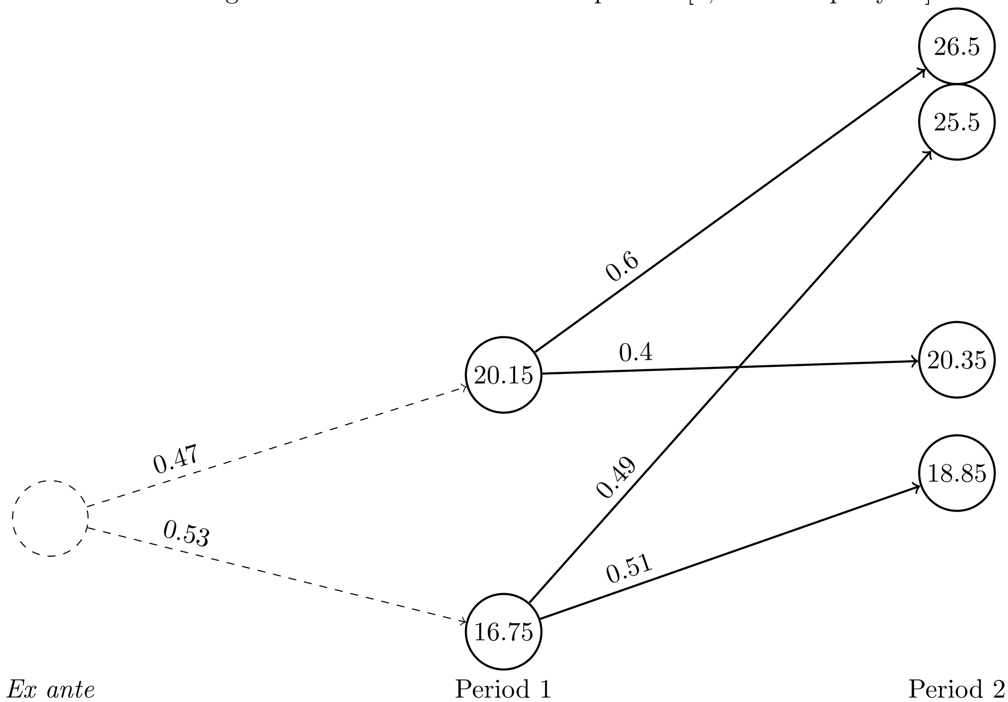
I use logarithmic utility from consumption and iso-elastic disutility from labor with compensated elasticity of 1:

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

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

[34] 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.

Figure 1: The estimated income process [1,000 EUR per year]



There are 7 parameters left to determine: the productivity at each history and the labor disutility parameter Γ. The productivities are pinned down with the first-order condition of labor supply

$$\theta = y(\theta)\sqrt{\Gamma \frac{1 - T(y(\theta))/y(\theta)}{1 - T'(y(\theta))}}, \tag{15}$$

where $T(y)$ is the actual Italian tax schedule.[35] Without the loss of generality Γ is set to 1 - by the first order condition (15) 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

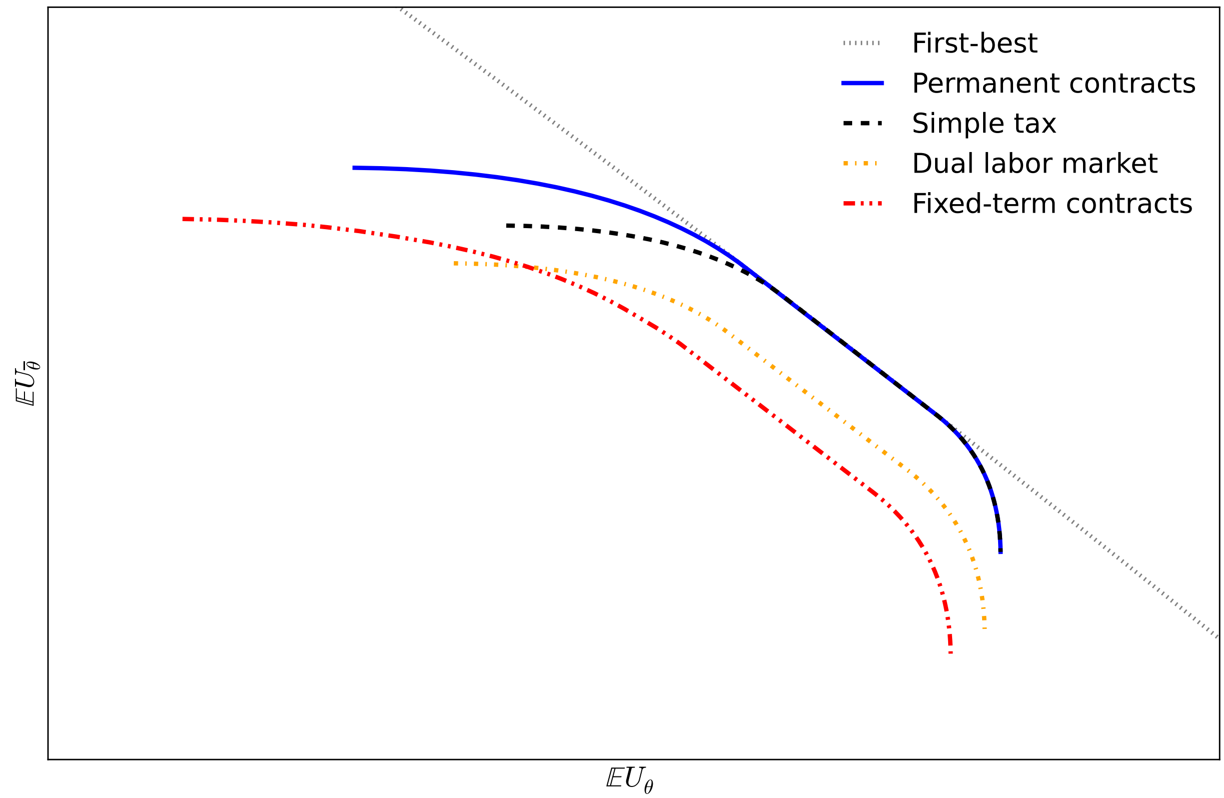
Figure 2 shows Pareto frontiers of four different regimes. 'Fixed-term contracts' regime corresponds to the NDPF economy, in which all workers receive fixed-term contracts and firms do not provide insurance. 'Dual labor market' frontier describes the economy in which the initially more productive type is employed permanently, while the other is employed on a fixed-term basis.[36] 'Permanent contracts' regime is characterized by both initial types receiving permanent employment. Finally, the 'Simple tax' regime corresponds to the restricted taxation problem considered in Section 5. In

[35] 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.

[36] I do not plot the Pareto frontier of the other configuration of contracts, in which the initially low productivity type receives permanent contract and the initially high type receives fixed-term contract. It is dominated by the permanent contracts case.

this regime all workers have permanent contracts and the allocation can be implemented with a simple consumption expenditure tax described by Theorem 4. I plot as well the Pareto frontier of the first-best allocation as an indicator of what is feasible if we abstract from 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.

Figure 2: 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 a regime in which at least one type of worker has permanent contract. Figure 2 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.

Table 4: Optimal allocations for different social welfare functions

Social preferences:	**utilitarian**	**libertarian**	**Rawlsian**	**anti-Rawlsian**
Welfare (cons. equiv.)				
Laissez-faire	100%	100%	100%	100%
Fixed-term contracts (NDPF)	102.8%	102.7%	107.2%	105.9%
Dual labor market	103.4%	103.3%	108.1%	104.8%
Simple tax	104.3%	104%	108.3%	105.8%
Permanent contracts (optimum)	104.3%	104%	108.3%	107.2%
Relative gain from permanent contracts	53.3%	49.3%	12.9%	20.3%

Note: The libertarian planner maximizes the average utility subject to no redistribution between the initial types. Taxation plays only an insurance role and workers would voluntarily decide to participate in a public insurance scheme. The anti-Rawlsian planner is the opposite of the Rawlsian planner and maximizes the utility of the most well-off type.

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 for the Rawlsian planner between the permanent contracts and the dual labor market regimes is 0.6% in consumption equivalent terms.

The simple tax allocation coincides with 'Permanent contracts' regime unless the planner strongly favors the initial high type. Labor distortions which are possible under permanent contracts, but not in the simple tax regime, are useful only when the planner wants to redistribute from the bottom to the top. In all the other cases, the constrained efficient allocations can be implemented with the consumption expenditure tax.

Table 4 compares regimes in terms of welfare under different social welfare functions. Utilitarian planner maximizes the expected utility of workers. Libertarian planner maximizes the expected utility of workers subject to the restriction of no transfers between initial types. The Rawlsian maximizes the utility of the least well-off worker, while the anti-Rawlsian planner cares about the most well-off worker. The benchmark allocation is a laissez-faire, which involve fixed-term contracts, no public insurance and uniform lump-sum taxation to cover the government expenditures. Allocations are compared using the consumption equivalent measure: by which factor we need to increase consumption of workers at each history in the laissez-faire allocation to obtain the same welfare as in the considered allocation. When we consider the less redistributive social planners (utilitarian and libertarian), the NDPF captures close to two-thirds of the gains from constrained efficient allocation. It means that the simpler tax that encourages firms' insurance improves upon the complicated tax system prescribed by NDPF by close to 50% in relative terms (the last row of Table 4). 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 corresponding Pareto frontiers are shown on Figure 3. Now the Rawlsian planner prefers the dual labor market regime. The welfare gain of the dual labor market over the permanent contract regime is significant: 1% in consumption equivalent terms.[37] Table 5 decomposes the welfare gain from a dual labor market into three components, corresponding to the change in consumption

[37] 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.

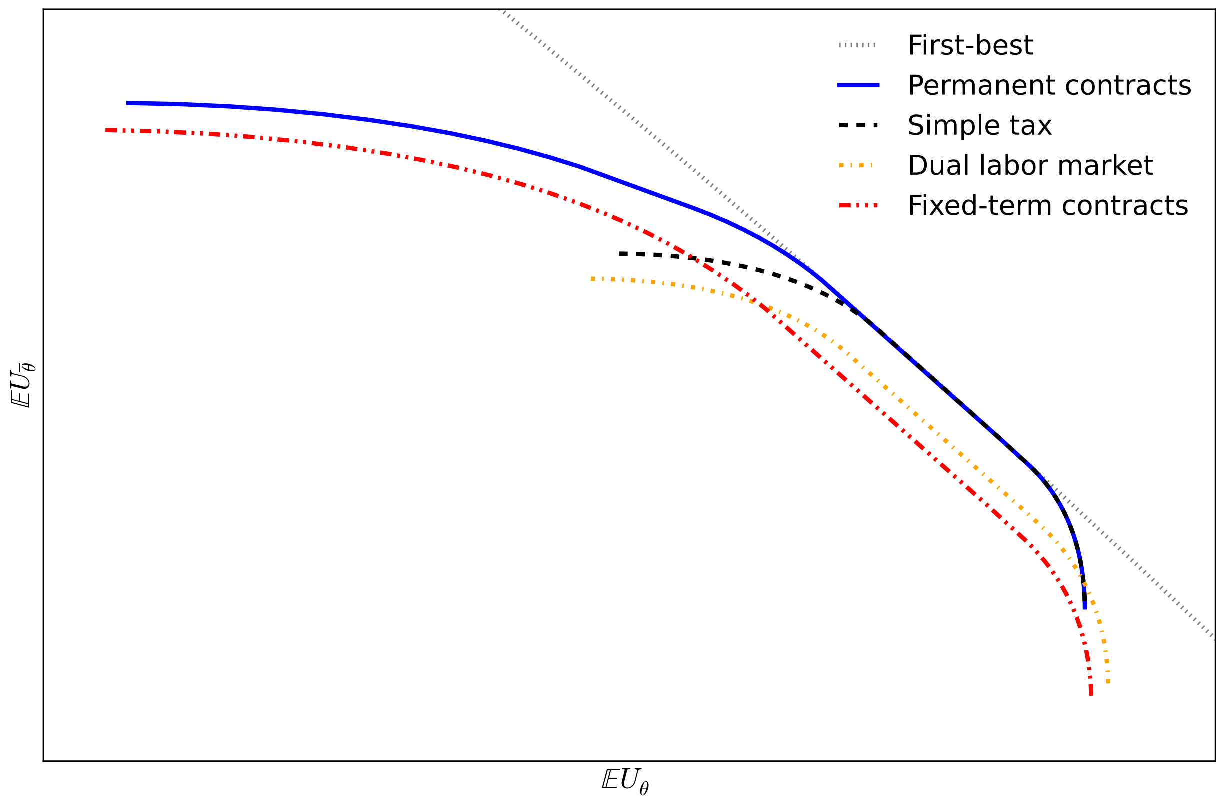
Table 5: Welfare impact of a dual labor market

	$\Delta^{insurance}$	$\Delta^{redistribution}$	$\Delta^{efficiency}$	**total change**
Original economy	-2.7%	2%	0.4%	-0.3%
Alternative economy	-2.8%	3%	0.7%	1%

Note: The decomposition is formally stated in Definition 8 in Appendix.

insurance, redistribution and efficiency. The main difference between the original economy and the alternative economy with increased initial differences is a greater gain in redistribution. The increased difference in initial productivities means that the deviating high type benefits more from income shifting: producing more in the initial period and getting paid in the second period. Income shifting is possible under permanent contracts and is prevented by the dual labor market.

Figure 3: 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 risks. 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 simple and realistic. The government optimally outsources insurance to firms by promoting permanent employment contracts. The only remaining role for the government is redistribution, which can be conducted with simple instruments. In a calibrated model of Italy all constrained efficient allocations which do not involve redistributing from the poor to the rich can be implemented with a comprehensible tax system: a time-invariant tax schedule that depends exclusively on current consumption expenditures.

Empowering the private sector to insure workers comes at a price. 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 levels of earnings. 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 heterogeneous production opportunities for firms would allow for analysis of temporary jobs, which in fact are the primary reason for existence of fixed-term contracts. Careful treatment of employees' outside option is vital to understanding insurance within firm. Specifically, the limited commitment of workers may be muted by search and mobility costs. Finally, in this paper there is no moral hazard problem within firms. It is worth examining how the optimal redistributive tax system interacts with the incentive provision in the private sector.

References

Ábrahám, Á., S. Koehne, and N. Pavoni (2016): "Optimal income taxation when asset taxation is limited," *Journal of Public Economics*, 136, 14–29.

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.

Bishow, J., and D. O. Parsons (2004): "Trends in severance pay coverage in the United States, 1980-2001," mimeo.

Bradford, D. F. (2000): *Taxation, wealth, and saving*. MIT Press.

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.

Diamond, P. A. (1998): “Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates,” *American Economic Review*, pp. 83–95.

Doligalski, P., and L. Rojas (2016): “Optimal Redistribution with a Shadow Economy,” *link*.

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 (2016): “Redistribution and Social Insurance,” *American Economic Review*, 106(2), 359–386.

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.

Hall, R. E., and A. Rabushka (1995): *The Flat Tax, 2nd ed.* Hoover Institution Press.

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).

Hopenhayn, H., and R. Rogerson (1993): "Job turnover and policy evaluation: A general equilibrium analysis," *Journal of Political Economy*, pp. 915–938.

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," *International Labour Review*.

Kreiner, C. T., S. Leth-Petersen, and P. E. Skov (2014): "Year-End Tax Planning of Top Management: Evidence from High-Frequency Payroll Data," *The American Economic Review*, 104(5), 154–158.

——— (2015): "Tax Reforms and Intertemporal Shifting of Wage Income: Evidence from Danish Monthly Payroll Records," *American Economic Journal: Economic Policy*, forthcoming.

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.

Lazear, E. P. (1990): "Job security provisions and employment," *The Quarterly Journal of Economics*, pp. 699–726.

LE MAIRE, D., AND B. SCHJERNING (2013): "Tax bunching, income shifting and self-employment," *Journal of Public Economics*, 107, 1–18.

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.

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

POSTEL-VINAY, F., AND H. TURON (2014): "The Impact of Firing Restrictions on Labour Market Equilibrium in the Presence of On-the-job Search," *The Economic Journal*, 124(575), 31–61.

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 Proofs from Section 2

Proof of Lemma 1. Take (σ, n) such that at some history $h \in \mathcal{H}$ the firm has positive expected profits. A competitor could profitably steal the worker by offering (σ', n'), where $\sigma' = \sigma$ and for all $r \in \mathcal{R}$ $n'_r(h) = n_r(h) - \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}\pi_h(c \circ r, n)/2\theta(h)$ and $n'_r(s) = n_r(s)$ for any other history $s \neq h$. This offer yields half of the profits of (σ, 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 fire the worker and incur the firing cost rather than to keep the worker. Hence, (σ, n) has to satisfy $\forall_{h\in\mathcal{H}} f(r(h)) \leq \sum_{r\in\mathcal{R}} \sigma(r)\mathbb{E}\pi_h(c \circ r, n) \leq 0$. $\square$

Proof of Lemma 2. The firm cannot expect losses initially, as she could offer a contract that yields zero profits by equalizing worker's output to worker's income at each history. Together with the limited commitment constraint at the initial type (3) it yields the zero profit condition (4).

Suppose there is an equilibrium (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 $\forall_{r\in\mathcal{R}}\forall_{h\in\mathcal{H}} n''_r(h) = n'_r(h) + \epsilon$, where $\epsilon > 0$. For epsilon sufficiently small (σ'', n'') yields positive profits and greater expected utility than (σ, n). Hence, (σ, n) cannot be an equilibrium.

Suppose that $(\sigma, n) \in \mathcal{E}(c, y, f)$ is not an equilibrium. It means that there is another (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 Proofs from Section 3

Lemma A.1. *Under full commitment on the labor market, all agents optimally enjoy full consumption insurance.*

Proof of Lemma A.1. Take any allocation (c, y, n) which is incentive-compatible (i.e. (n, σ^*) is the equilibrium, given (c, y)) and do not involve full consumption insurance. For each type $h \in \mathcal{H}$ find a full-insurance consumption level $\bar{c}(h)$ with equality $\sum_{t=1}^{\bar{t}} \beta^{t-1} u(\bar{c}(h)) = \sum_{s\in\mathcal{H}(h_1)} \beta^{1-|s|} \mu(s \mid h_1) u(c(s))$. Set $\bar{y}(h)$ to the average of histories of this length: $\sum_{s\in\mathcal{H}_{|h|}(h_1)} \mu(s \mid h_1) y(s)$. This way both the lifetime consumption and labor income are deterministic functions of the initial type report. As the worker receives more consumption insurance, the planner frees some resources.

Before we prove that truth-telling is the equilibrium strategy given the new outcomes, lets define a useful class of reporting strategies. Take some pure reporting strategy $r \in \mathcal{R}$. The *statistical mimicking* strategy σ_r^{stat} is a mixed reporting strategy such that

$$\forall_{s_1\in\Theta_1}\forall_{h\in\mathcal{H}}\ \mu(h \mid r(s_1)) = \sum_{r'\in\mathcal{R}} \sigma_r^{stat}(r') \sum_{h'\in\mathcal{H}} \mu(h' \mid s_1)\mathbb{I}_{r'(h')=h},$$

where $\mathbb{I}$ is an indicator function. Statistical mimicking strategy generates the distribution of type reports of some initial type s_1 consistent with truthful reporting of initial type $r(s_1)$. Note that the expected lifetime labor income of type s_1 from following σ_r^{stat} is equal to the expected lifetime income of initial type $r(s_1)$.

In order to check that the truthful reporting is the equilibrium strategy given $(\bar{c}, \bar{y})$, consider any pure reporting strategy $\bar{r} \in \mathcal{R}$. The utility of any initial type from following this strategy given outcomes $(\bar{c}, \bar{y})$ is equal to the utility from following the strategy $\sigma_{\bar{r}}^{stat}$ with outcomes (c, y). To see this, note that the utility from consumption and the expected lifetime income generated by $\sigma_{\bar{r}}^{stat}$ given (c, y) are identical to those generated by reporting $\bar{r}$ given $(\bar{c}, \bar{y})$. Since lifetime incomes are the same in the two cases, the labor supply allocation will be the same as well. Therefore, the utility from following $\bar{r}$ given $(\bar{c}, \bar{y})$ is equal to the utility from following $\sigma_{\bar{r}}^{stat}$ given the original mechanism (c, y). Since in the original mechanism the expected utility from any reporting strategy was bounded above by the truthful reporting, the expected utility from following $\bar{r}$ is also weakly lower then the expected utility from truth-telling in the new mechanism $(\bar{c}, \bar{y})$. What remains to be shown is that incentive compatibility holds also for the mixed reporting strategies. For any $\bar{\sigma} \in \Delta_{\mathcal{R}}$ we can define another mixed reporting strategy σ' such that $\sigma'(r') = \sum_{r\in\mathcal{R}} \bar{\sigma}(r)\sigma_r^{stat}(r')$. By the argument above, the expected utility from following $\bar{\sigma}$ given $(\bar{c}, \bar{y})$ is equal to the expected utility from following σ' given (c, y). Since the original mechanism implements truth-telling, so does the new mechanism. □

Proof of Theorem 1. First, Lemma A.1 shows that it is always optimal to implement full consumption insurance allocation, in which the lifetime consumption and the lifetime labor income is determined by the initial type report. It means that we can represent the expected utility of the initial type h_1 that reports type s_1 with $V_{h_1}(C(s_1), Y(s_1))$. Below I show first that at the optimum only the incentive-compatibility constraints with respect to mixed strategies can bind. Secondly, I prove that (9) is a relevant incentive-compatibility constraint with respect to mixed reporting strategies.

First, I will show that only the incentive compatibility constraints corresponding to mixed reporting strategies can be binding. Suppose that there is some pure reporting strategy $r \neq r^*$ such that $V_s(C(s), Y(s)) = V_s(C(r(s)), Y(r(s)))$. Consider reporting strategy σ which mixes between the truthful reporting and r: $\sigma(r^*) + \sigma(r) = 1$, $\sigma(r) \in (0, 1)$. Denote $\bar{Y} \equiv \sigma(r^*)Y(s) + \sigma(r)Y(r(s))$ and take a particular labor allocation $\bar{n}$ defined as $\forall_{s'\in\mathcal{H}(s)} \bar{n}(s') = \sigma(r^*)n^{FC}_{Y(s)}(s) + \sigma(r)n^{FC}_{Y(r(s))}(s)$ which generates $\bar{Y}$. Since v is strictly convex, $n^{FC}_{\bar{Y}}(s') \neq \bar{n}(s')$, so this labor allocation is suboptimal. Nevertheless, as we will see, the worker prefers to deviate to σ even with a suboptimal labor allocation after deviation. Let's compare the worker's payoff from σ with the payoff from truthful reporting:

$$
\begin{aligned}
&\sum_{r'\in\mathcal{R}} \sigma(r')V_s(C(r'(s), \bar{Y}) - V_s(C(s), Y(s)) \\
&= \sum_{r'\in\mathcal{R}} \sigma(r') \left(V_s(C(r'(s), \bar{Y}) - V_s(C(r'(s), Y(r'(s)))\right) \\
&= \sum_{s'\in\mathcal{H}(s)} \beta^{|s'|}\mu(s' \mid s) \left(\sum_{r'\in\mathcal{R}} \sigma(r')v\left(n^{FC}_{Y(r(s))}(s')\right) - v(n^{FC}_{\bar{Y}}(s')) \right) \\
&\geq \sum_{s'\in\mathcal{H}(s)} \beta^{|s'|}\mu(s' \mid s) \left(\sum_{r'\in\mathcal{R}} \sigma(r')v\left(n^{FC}_{Y(r(s))}(s')\right) - v\left(\bar{n}(s')\right) \right) > 0.
\end{aligned}
$$

The first equality comes from the fact that the pure reporting strategy r provides as much utility to initial type s as truthful revelation. Then we can cancel out the utility from consumption, which leads to the second equality. The first inequality is implied by the fact that $\bar{n}$ is not a utility maximizing choice of labor supply that generates $\bar{Y}$. The final result comes from Jensen's inequality, since v is strictly concave. We have seen that whenever any pure reporting strategy gives as much expected utility as the truthful reporting, the agent would be strictly better of by mixing between the two. Therefore, only the incentive constraints with respect to mixed strategies can bind in the optimum.

Consider the choice of the mixed reporting strategy of some initial type s given the schedules of lifetime consumption C and lifetime income Y:

$$\max_{\sigma \in \Delta_{\mathcal{R}}} \sum_{r \in \mathcal{R}} \sigma(r) V_s \left(C(r(s)), \sum_{r \in \mathcal{R}} \sigma(r) Y(r(s)) \right).$$

The derivative of the objective function with respect to $\sigma(r)$, under normalization $\sigma(r^*) = 1 - \sum_{r \neq r^*} \sigma(r)$, is given by

$$\bar{\beta} u(C(r(s))/\bar{\beta}) - \bar{\beta} u(C(s)/\bar{\beta}) - (Y(r(s))) - Y(s))\, \phi_s \left(\sum_{r \in \mathcal{R}} \sigma(r) Y(r(s)) \right),$$

where $\phi_s(Y)$, defined by (8), equals $-\partial V_s(C,Y)/\partial Y$. This problem is concave, since $\phi_s'(Y) > 0$. Hence, the necessary and sufficient condition for truth-telling is that the above derivative is non-positive when evaluated at the truthful reporting

$$\forall_{s' \in \Theta_1} \bar{\beta} u(C(s')/\bar{\beta}) - \bar{\beta} u(C(s)/\bar{\beta}) - (Y(s') - Y(s))\, \phi_s\left(Y(s)\right) \leq 0.$$

The rearrangement of terms generates the incentive-compatibility conditions (9). □

Proof of Proposition 1. Take $s, s' \in \Theta_1$ such that $Y(s) > Y(s')$. Then incentive compatibility constraints (9) preventing s from mimicking s' and *vice versa* imply that

$$\phi_{s'}(Y(s')) \geq \frac{\bar{\beta}\left(u(C(s)/\bar{\beta}) - u(C(s')/\bar{\beta}) \right)}{Y(s) - Y(s')} \geq \phi_s(Y(s)).$$

Finally note that $\phi_s(Y(s)) = (1 - T'(Y(s)))u'(C(s)/\bar{\beta})$. □

Proof of Proposition 2. Set up a Lagrangian $\mathcal{L}$ corresponding to the maximization problem in Theorem 1. Denote the top type by $\overline{\theta}$. Denote the multiplier w.r.t. the resource constraint by η. We know that in the optimum some downwards incentive constraints of the top type will bind. Moreover, no lower type is tempted to mimic the top type - otherwise, the planner could assign them the top type's income and consumption and get additional resources, since they would end up paying higher taxes. Denote the multipliers with respect to the incentive constraints preventing the top type from mimicking some type θ by ξ_θ. The derivatives of the Lagrangian w.r.t. $C(\overline{\theta})$ and $Y(\overline{\theta})$ are

$$\begin{aligned}
\frac{\partial \mathcal{L}}{\partial C(\overline{\theta})} =& \mu(\overline{\theta}) u'(C(\overline{\theta})) - \mu(\overline{\theta})\eta + \sum_{\theta \in \Theta_1} \xi_\theta u'(C(\overline{\theta})), \\
\frac{\partial \mathcal{L}}{\partial Y(\overline{\theta})} =& -\mu(\overline{\theta}) \phi_{\overline{\theta}}(Y(\overline{\theta})) + \mu(\overline{\theta})\eta - \sum_{\theta \in \Theta_1} \xi_\theta \left(\phi_{\overline{\theta}}(Y(\overline{\theta})) + \left(Y(\overline{\theta}) - Y(\theta)\right) \phi_{\overline{\theta}}'(Y(\overline{\theta})) \right).
\end{aligned}$$

By setting the derivatives to zero and combining the two equations we get

$$\frac{\phi_{\overline{\theta}}(Y(\overline{\theta}))}{u'(C(\overline{\theta}))} = \frac{\mu(\overline{\theta}) + \sum_{\theta \in \Theta_1} \xi_\theta}{\mu(\overline{\theta}) + \sum_{\theta \in \Theta_1} \xi_\theta \left(1 + \left(Y(\overline{\theta}) - Y(\theta)\right) \frac{\phi'_{\overline{\theta}}(Y(\overline{\theta}))}{\phi_{\overline{\theta}}(Y(\overline{\theta}))}\right)}.$$

The term $\left(Y(\overline{\theta}) - Y(\theta)\right) \frac{\phi'_{\overline{\theta}}(Y(\overline{\theta}))}{\phi_{\overline{\theta}}(Y(\overline{\theta}))}$ is positive for all $\theta < \overline{\theta}$, hence the ratio above is smaller than 1. Consequently, the labor supply of the top type is distorted downwards. □

A.3 Proofs from Section 4

Proof of Lemma 3. Take any (σ, n_σ) which is consistent with limited commitment constraints (3) and yields non-negative profits ex ante. Neither workers nor firms can have incentives to terminate the contract (i.e. condition (3) is satisfied) for any pure reporting strategy r such that $\sigma(r) > 0$. It is possible only if (σ, n_σ) yields zero profits at each pure reporting strategy that can be drawn with a positive probability. Otherwise there is some pure strategy with a positive probability which yields positive profits, which violates (3). Furthermore, $\sum_{r \in \mathcal{R}} \sigma(r) \mathbb{E}U(c \circ r, n_r) \leq \mathbb{E}U(c \circ r', n_{r'})$ for some r' such that $\sigma(r') > 0$. Therefore, $(r', n_{r'})$ yields the same profits and weakly greater utility than (σ, n_σ). □

Proof of Lemma 4. 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 $(\sigma^*, n^{FC}) \in \mathcal{E}(c^{FC}, y, f)$, where only permanent contracts are used: $\forall_{h \in \mathcal{H}} f(h) = \bar{f}$. For any 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 one, then (σ^*, n^{FC}) would not be an equilibrium contract under full commitment. □

Lemma A.2. *Suppose that (i) there are two time periods, (ii) there are two productivity levels that are independent over time: $\Theta_1 = \Theta_2 = \{0, 1\}$. Consider allocation (c, y, n) implemented with* *mechanism (c, y, f), where $f(1) = 0$ and c is consistent with the inverse Euler equation. The planner can implement the same allocation with a mechanism (c, y, f'), where $f'(1) = \bar{f}$ and $f'(0) = f(0)$.*

Proof of Lemma A.2. First, note that because of zero productivity the initial low type is unable to mimic the initial high type. Second, I will show that the firm would choose the same labor supply allocation of the initial high type under the new contract assignment. The inverse Euler equation together with fixed term contract implies that $u'(c(1,1)) > u'(c(1)) > u'(c(1,0))$. The labor supply $n(1)$ is undistorted since the initial incentive compatibility constraint preventing type 0 from mimicking type 1 is slack. What is more, $n(1,1)$ is undistorted as well, since insurance against the second period productivity risk requires that type $(1,1)$ finances consumption of the type $(1,0)$ ('no distortion at the top'). Hence, we have

$$v'(n(1)) = u'(c(1)) > u'(c(1,1)) = v'(n(1,1)).$$

It means that, when the high type has permanent contract, the firm cannot improve his allocation of labor by shifting labor supply from the history $(1,1)$ to the initial history. Of course, the firm cannot shift labor from the other second period history, since $n(1,0) = 0$. Therefore, the labor supply allocation of the high type would remain unchanged after the increase in the firing cost. □

Proof of Theorem 3. Consider the mechanism (c, y, f) with the equilibrium (r^*, n) where the top taxpayer $\theta \in \Theta_1$ has a fixed term contract: $f(\theta) = 0$. Consider a new mechanism (c', y', f') in which θ has permanent contract, the full consumption insurance and the same level of utility as before. More insurance means that the planner saves some resources. If any other type $\theta' \in \Theta_1$ wants to mimic θ, assign $(c'(\theta'), y'(\theta'), f'(\theta')) = (c'(\theta), y'(\theta), f'(\theta))$. They are weakly better-off as well. Furthermore, the planner has more resources - the mimicking types become top-taxpayers and pay higher taxes.

□

Definition 8 (Welfare impact decomposition)**.** Consumption allocation c_1 is defined as follows. c_1 is equal to c for all $h \in \mathcal{H} \backslash \mathcal{H}(\underline{\theta})$. For the histories following from the initial type $\underline{\theta}$, it is defined as

$$c_{1|\mathcal{H}(\underline{\theta})} \in \arg \max_{\tilde{c}: \mathcal{H}(\underline{\theta}) \to \mathbb{R}_+} \mathbb{E}U_{\underline{\theta}}(\tilde{c}, n)$$

subject to keeping the present value of consumption constant $\sum_{h \in \mathcal{H}(\underline{\theta})} R^{-|h|} \mu(h)(\tilde{c}(h) - c(h)) = 0$ and the incentive constraints corresponding to insurance

$$\forall_{\substack{r \in \mathcal{H}(\underline{\theta}) \to \mathcal{H}(\underline{\theta}) \\ \text{s.t.} \exists_{r' \in \mathcal{R}} \forall_{h \in \mathcal{H}(\underline{\theta})} r(h) = r'(h)}} \quad \mathbb{E}U_{\underline{\theta}}(\tilde{c}, n) \geq \mathbb{E}U_{\underline{\theta}}(\tilde{c} \circ r, \bar{n}(r)).$$

Consumption allocation c_2 (and the associated allocation of income y_2) is defined by

$$(c_2, y_2) \in \arg \max_{\substack{\tilde{c} : \mathcal{H} \to \mathbb{R}_+ \\ \tilde{y} : \mathcal{H} \to \mathbb{R}}} W(\tilde{c}, n)$$

subject to the budget constraint $\sum_{h\in\mathcal{H}} R^{-|h|}\mu(h)(\tilde{c}(h)-c(h))=0$ and the incentive compatibility constraints

$$\forall_{r\in\mathcal{R}}\forall_{n_r:\mathcal{H}\to\mathbb{R}_+}\mathbb{E}U_{\underline{\theta}}(\tilde{c},n)\geq\mathbb{E}U_{\underline{\theta}}(\tilde{c}\circ r,n_r),$$

where n_r satisfy zero profit (4) and limited commitment (3) constraints given the labor income allocation $\tilde{y}$. Finally, consumption allocation c' (and the associated allocation of income y') is defined by

$$(c',y')\in\arg\max_{\substack{\tilde{c}\,:\,\mathcal{H}\to\mathbb{R}_+\\ \tilde{y}\,:\,\mathcal{H}\to\mathbb{R}}} W(\tilde{c},\tilde{n})$$

subject to the budget constraint $\sum_{h\in\mathcal{H}} R^{-|h|}\mu(h)(\tilde{y}(h)-\tilde{c}(h))=0$ and the equilibrium constraint

$$(r^*,\tilde{n})\in\mathcal{E}(c',y',f').$$

The labor supply allocation n' is the labor allocation corresponding to the truthful reporting strategy in $\mathcal{E}(c',y',f')$.

Proof of Lemma 5. $\Delta^{insurance}$ is non-positive, since both c_1 and c have the same net present value, but c_1 has to satisfy additional incentive-compatibility constraints implied by fixed-term contract. Specifically, full consumption insurance of $\underline{\theta}$ is ruled out. $\Delta^{efficiency}$ is non-negative, since the planner can choose outcomes (c_2,y_2,f'). Under these outcomes, the welfare is weakly higher than $W(c_2,n_2)$, as the firms can optimize over labor. Note that when we found the allocation c_2, the firms were allowed to optimize over labor only when deviating.

Suppose that $\underline{\theta}\in\arg\max_{\theta\in\Theta_1}\lambda(\theta)u'(c_1(\theta))$. It means that the planner wants to redistribute to $\underline{\theta}$. By assigning fixed-term contract to this type, the incentive constraint that prevent redistribution from other initial types to $\underline{\theta}$ are (weakly) relaxed for two reasons: the consumption of $\underline{\theta}$ is more volatile and the fixed-term contract prevents labor smoothing after deviation. Since incentive constraints are relaxed, the planner can redistribute to $\underline{\theta}$, raising the social welfare ($\Delta^{redistribution}\geq 0$). In the second case, when $\underline{\theta}\in\arg\min_{\theta\in\Theta_1}\lambda(\theta)u'(c_1(\theta))$, the planner wants to redistribute *from* $\underline{\theta}$. However, assigning fixed-term contract weakly decreases utility of this type because of more volatile consumption. Hence, $\underline{\theta}$ is even more tempted to mimic other types and the redistribution has to be reduced, leading to lower welfare ($\Delta^{redistribution}\leq 0$). □

Definition 9. A *one-shot deviation* $d_{h,h'}$ is a reporting strategy such that *(i)* $\forall_{s\in\mathcal{H}/\mathcal{H}(h)}d_{h,h'}(s)=s$ and *(ii)* $\forall_{s\in\mathcal{H}(h)}d_{h,h'}(s)=(h',s_{|h|+1},...,s_{|s|})$.

Lemma A.3. *Suppose that only fixed-term contracts are used. Under Assumption 1, if the incentive constraints with respect to one-shot deviations hold, the incentive constraints with respect to all reporting strategies are satisfied.*

Proof of Lemma A.3. First I'll prove a useful property that holds under Assumption 1. Take any reporting strategy r' and some history h. Construct another reporting strategy r'' such that $r''(s)=s$ for $s\notin\mathcal{H}(r'(h))$ and $r''(s)=r'(h,s_{|h|+1},...,s_{|s|})$ for $s\in\mathcal{H}(r'(h))$. For any $\theta\in\Theta_{|h|+1}$

$$\mathbb{E}U_{(h,\theta)}(c \circ r', \bar{n}(r')) = \mathbb{E}U_{(r'(h),\theta)}(c \circ r'', \bar{n}(r'')). \tag{16}$$

The full support assumption guarantees that $(r'(h), \theta)$ is a history with a positive probability. Note that history of reports on both sides of equation are identical. Moreover, the last productivity draw is the same, so by the Markov property productivity distribution is the same. Finally, the function from new productivity draws to reports is identical as well. Hence, the payoffs are equal.

Take any allocation (c, y, n). I'll show that if (i) the payoff from any reporting strategy r that is truthful before history $h \in \mathcal{H}_{t+1}$ is dominated by the payoff from the one-shot deviation $d_{h,r(h)}$ and (ii) the incentive constraints w.r.t. all one-shot deviations that are truthful before period $t+1$ hold, then the payoff from any reporting strategy r' that is truthful before any history $h' \in \mathcal{H}_t$ is dominated by the payoff from the one-shot deviation $d_{h',r'(h')}$.

Take some reporting strategy r' that is truthful before history $h \in \mathcal{H}_t$ and define r'' as in the first paragraph of this proof. For any $\theta \in \Theta_{t+1}$

$$\begin{aligned}
\mathbb{E}U_{(h,\theta)}(c \circ r', \bar{n}(r')) &= \mathbb{E}U_{(r'(h),\theta)}(c \circ r'', \bar{n}(r'')) \\
&\leq \mathbb{E}U_{(r'(h),\theta)}\left(c \circ d_{(r'(h),\theta),r''(r'(h),\theta)}, \bar{n}\left(d_{(r'(h),\theta),r''(r'(h),\theta)}\right)\right) \\
&\leq \mathbb{E}U_{(r'(h),\theta)}(c, n) \\
&= \mathbb{E}U_{(h,\theta)}\left(c \circ d_{h,r'(h)}, \bar{n}\left(d_{h,r'(h)}\right)\right).
\end{aligned}$$

The first equality comes from the property (16). The second step is implied by the assumption (i) that the one-shot deviations dominate all other reporting strategies that are truthful before period $t+1$. The consecutive inequality is implied by incentive compatibility w.r.t. one-shot deviations. The final equation is again implied by (16). Summing up payoffs for all $\theta \in \Theta_{t+1}$, weighted by their conditional probability, and adding the instantaneous utility at the history h leads to

$$\mathbb{E}U_h(c \circ r', \bar{n}(r')) \leq \mathbb{E}U_h\left(c \circ d_{h,r'(h)}, \bar{n}\left(d_{h,r'(h)}\right)\right). \tag{17}$$

This inequality means that the payoff from r' is bounded above by the payoff from the corresponding one-shot deviation, which concludes the proof of the induction step. Finally, note that at the terminal period $\bar{t}$ the only reporting strategies that are truthful before the terminal period are one-shot deviations. Hence, by induction, if incentive constraints with respect to all one-shot deviations are satisfied, the incentive constraints w.r.t. all possible reporting strategies are satisfied. □

Proof of Proposition 3. 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 θ under truthful revelation. Denote by n_0 the labor supply allocation when type $\underline{\theta}$ has permanent contract. When $\underline{\theta}$ has a permanent contract, the incentive constraint that prevents the redistribution from type θ to $\underline{\theta}$ is

$$C(\theta) - \sum_{s \in \mathcal{H}(\theta)} \beta^{|s|-1} \mu(s \mid \theta) v(n(s)) \geq C(\underline{\theta}) - \sum_{s \in \mathcal{H}(\theta)} \beta^{|s|-1} \mu(s \mid \theta) v(\tilde{n}(s)), \tag{18}$$

where $\tilde{n}$ produces $Y(\underline{\theta})$ and satisfies the labor smoothing condition (8) whenever it is consistent with limited commitment. When $\underline{\theta}$ has a fixed-term contract, keeping the net present value of consumption and the allocation of labor unchanged, the analogous incentive constraint is

$$C(\theta) - \sum_{s \in \mathcal{H}(\theta)} \beta^{|s|-1} \mu(s \mid \theta) v(n(s)) \geq C(\underline{\theta}) - \sum_{s \in \mathcal{H}(\theta)} \beta^{|s|-1} \mu(s \mid \theta) v(\hat{n}(s)), \tag{19}$$

where $\hat{n}(s) = \frac{r_{\theta,\underline{\theta}}(s)}{\theta(s)} n_0(s)$. First, by Lemma A.3 we can focus only on one-shot deviation. Second, since future productivities are independent of initial draw, the net present value of consumption of deviating type θ is equal to $C(\underline{\theta})$ and $Y(\underline{\theta})$. I will show that the right-hand side of (18) is strictly greater than of (19). Note that $\tilde{n}$ and $\hat{n}$ produce the same output $Y(\underline{\theta})$. Moreover, the labor supply allocation n satisfies

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

Since $\theta > \underline{\theta}$, this implies that

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

The initial type θ could reduce the disutility from labor by producing more in the initial period, which does not violate the limited commitment constraints. It means that $\hat{n}(\theta) < \tilde{n}(\theta)$ and hence the right-hand side of (18) is strictly greater than of (19). Hence, the incentive constraint is relaxed when $\underline{\theta}$ receives fixed-term contract. Since it is true for all types $\theta \in \Theta \backslash \underline{\theta}$, we have $\Delta^{redistribution} > 0$. □

Proof of Proposition 4. Note that assigning fixed-term contract to $\underline{\theta}$, while keeping outcome functions c and y constant, does not change the utility from truthfully or untruthfully reporting $\underline{\theta}$ ($\Delta^{insurance} = \Delta^{redistribution} = 0$). Below I show that when $\underline{\theta}$ has fixed-term contract, the planner can perturb the outcome functions of this type to save resources without changing his utility level nor violating any incentive constraint.

Take history $h \in \mathcal{H}(\underline{\theta})$ with $\theta(h) = \max \Theta_{|h|}$ at which the labor supply is distorted downwards: $\theta(h)u'(c(h) > v'(y(s)/\theta(s))$. Perturb consumption and income such that $\theta(h)u'(c'(h)) = v'(y'(h)/\theta(h))$ and the instantaneous utility at this history is unchanged. Since distortions are lifted, the planner obtains additional resources. Furthermore, since h is the most productive type, lifting the downward distortion relaxes the incentive constraints w.r.p. other types $h' \in \mathcal{H}(h^{-1})$. The incentive constraints corresponding to deviations at earlier dates are unaffected by Assumption 1 and Lemma A.3. The utility from a one shot deviation $\mathbb{E}U_{(s,\theta)}(c \circ d_{s,h^{-1}}, n_{d_{s,h^{-1}}}) = \mathbb{E}U_h(c, n)$ is unaffected by the perturbation for any $\theta \in \Theta_{|h|}$. Hence, when $\underline{\theta}$ has fixed-term contract, the planner can lift some future distortions and obtain additional resources without violating incentive compatibility. These resources can be spend on uniform raise of expected utility of all types, leading to $\Delta^{efficiency} > 0$. □

A.4 Proofs from Section 5

Lemma A.4. *Under Assumption 2 the function $V_\theta(C, Y)$ satisfies the Spence-Mirrlees single-crossing condition.*

Proof of Lemma A.4. The Spence-Mirrlees condition states that $-\frac{\partial V_\theta(C,Y)}{\partial Y}\left(\frac{\partial V_\theta(C,Y)}{\partial C}\right)^{-1}$ is non-increasing with θ. Since the denominator is positive and constant in θ, it is enough to show that $\frac{\partial V_\theta(C,Y)}{\partial Y}$ is non-decreasing with θ.

Note that $\frac{\partial V_\theta(C,Y)}{\partial Y} = -\frac{v'(n(h))}{\theta(h)}$ for any $h \in \mathcal{H}(\theta)$. If we fix Y, the labor supply depends only on the initial type h and current productivity θ, so we can write it as $n_h(\theta)$. Note that $n_h(\theta)$ is increasing in θ. Let's define the extension of $n_{h_1}(\theta)$ to all possible productivity realizations with a step function $\bar{n}_h(\theta) \equiv \max_{\theta' \in \Theta(h)\cap[0,\theta]} n_h(\theta')$, where $\Theta(h) \equiv \{\theta(s) : s \in \mathcal{H}(h)\}$ is a set of all possible productivity realizations following the history h.

Suppose that $\frac{\partial V_{h_1}(C,Y)}{\partial Y} > \frac{\partial V_{s_1}(C,Y)}{\partial Y}$ for some initial types $s_1 > h_1$. It means that for any productivity level $\theta \in \Theta(h_1) \cup \Theta(s_1)$ we have $n_{s_1}(\theta) > n_{h_1}(\theta)$. Now we have

$$\sum_{\theta\in\mathbb{R}_+} \hat{\mu}(\theta \mid h_1)\theta n_{h_1}(\theta) = \sum_{\theta\in\mathbb{R}_+} \hat{\mu}(\theta \mid h_1)\theta \bar{n}_{h_1}(\theta) \leq_{FOSD} \sum_{\theta\in\mathbb{R}_+} \hat{\mu}(\theta \mid s_1)\theta \bar{n}_{h_1}(\theta) < \sum_{\theta\in\mathbb{R}_+} \hat{\mu}(\theta \mid s_1)\theta n_{s_1}(\theta).$$

The weak inequality is implied by the first-order stochastic dominance (Assumption 2), since $\bar{n}_{h_1}(\theta)$ is a non-decreasing function of θ. The second inequality comes from $n_{s_1}(\theta) > n_{h_1}(\theta') = \bar{n}_{h_1}(\theta)$ for some $\theta' \leq \theta$. The left-hand side is the lifetime income of initial type h_1 divided by $\sum_{t=1}^{\bar{t}}$, while the right-hand side is the lifetime income of s_1. Since we assumed that their lifetime incomes are both equal to Y, we have a contradiction. Therefore $\frac{\partial V_\theta(C,Y)}{\partial Y}$ is non-decreasing in θ. ☐

Proof of Proposition 5 . 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 2 and Lemma A.4. Hence, what remains to be shown are the relevant elasticities, which I derive below.

First, recall the definition of $\phi_\theta(Y)$ in (8). Let's define the marginal tax rate $T'(Y(\theta))$ as $1 - \frac{\phi_\theta(Y(\theta))}{u'\left(C(\theta)/\bar{\beta}\right)}$. I will derive the elasticities by varying the marginal tax rate. The compensated elasticity is given by

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

We can derive ϕ' with the implicit function theorem. First, we can use (8) express $n(h)$ as $g(\theta(h)\,\phi_{h_1}(Y(h_1)))$, where g is an inverse function of v'. Plug this expression into the zero profit condition to get

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

By the implicit function theorem we have

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

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

Hence, we have

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

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, weighted by the realized output. The uncompensated elasticity is given by

$$\bar{\zeta}^u = \frac{\partial Y\left(\theta\right)}{\partial 1 - T'\left(\theta\right)} \frac{1 - T'\left(\theta\right)}{Y\left(\theta\right)} = \bar{\zeta}^u + \frac{u''\left(C(\theta)/\bar{\beta}\right)}{\bar{\beta}\phi'_\theta\left(Y\left(\theta\right)\right)}\left(1 - T'\left(\theta\right)\right).$$

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

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

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. □

Proof of Lemma 6. It follows from Lemma 2, with individual consumption determined by the usual budget constraint. □

Proof of Lemma 7. We can apply the proof of Theorem 2. □

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 $\tilde{\mathcal{E}}(T)$. 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 does 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 $\bar{T}_x(x_t) + \alpha(x_t - x_0)^2$ is convex. Then the reasoning above applies. $\square$

Proof of Proposition 6. I will show that the inverse Euler equation holds. The proof follows 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

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

and c' 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). $\square$

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