Social Preferences, Skill Segregation and Wage Dynamics*

Antonio Cabrales† Antoni Calvó-Armengol‡ Nicola Pavoni§

Abstract

We study the earning structure and the equilibrium assignment of workers to firms in a model in which workers have social preferences and skills are perfectly substitutable in production. We allow firms to offer long terms contracts and for frictions in the labour market in the form of mobility costs. For low moving costs between firms, heterogeneous productivities lead to widespread workplace skill segregation and the whole market wage dispersion is explained by between firms differences. In a labor market with intermediate levels of mobility costs, segregation is more moderate and wage dispersion arises both within and across firms. For high levels of moving costs the whole wage dispersion is within the firm, and becomes zero when the moving costs are sufficiently high. We show that long terms contracts in the presence of social preferences associate within-firm wage dispersion with novel 'internal labor market' features such as a dynamic form of wage compression, gradual promotions, and wage non-monotonicity.

*We gratefully acknowledge the financial help from Barcelona Economics CREA, Fundación BBVA, and from the Spanish Ministry of Science and Technology under grants SEC2003-03474 and BEC2002-2130. The usual disclaimer applies.

†Departament d'Economia i Empresa, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, 08005 Barcelona, Spain. Email: antonio.cabrales@upf.edu. http://www.econ.upf.edu/~cabrales

‡Universitat Autònoma de Barcelona and CEPR. Email: antoni.calvo@uab.es. http://selene.uab.es/acalvo

§University College of London and Institute of Fiscal Studies, London. Email: n.pavoni@ucl.ac.uk. http://www.ucl.ac.uk/~uctpnpa/

1 Introduction

The upward trend in inequality that has taken place in the United States and other countries since the 70s has generated a renewed interest in problems associated with earnings inequality. The trend towards allocating high and low skilled workers into separate firms is perceived as a key element to understand earnings dispersion.

In this paper we study phenomena such as workers' flows, skill segregation, within and between firms wage dispersion as market equilibrium outcomes in environments with no skill complementarities in production. Instead, we assume that preferences of workers depend not only on their compensations, but also on that of their co-workers. This assumption is consistent with a wide body of evidence showing that preferences of individuals between allocations do not depend only on their own material well-being. Rather, the actions and material allocations of other individuals impact directly a person's utility, and are thus taken into account when making a decision.

We consider a labor market in which risk-neutral firms compete for risk-averse workers of heterogeneous quality. The efficiency units of workers' labor are perfect substitutes. That is, some workers are more productive/skilled than others, but workers of different skills are perfectly substitutable in some fixed proportions. Firms compete by offering long-term contracts. The firms can commit to the contracts, but the workers can always accept external offers.[1] The quality of the workers is not perfectly observable ex-ante but their performance over time slowly reveals (with some noise) this quality. The workers have "social preferences," that is, their final utility is affected by that of others. But which others? Most standard models of social preferences focus on two or three person games between an employer and one or two employees. But in a model of a market, the range of interpersonal comparisons of utility is an important consideration. We assume that these comparisons do not span the whole population, but only individuals who work in the same firm, and have similar career histories within the firm. For any given worker and period, we call his *reference group* the set of individuals over whom his social preferences' comparisons take place in that period.

With the structure of our model, and the traditional "selfish" preferences, the equilibria would not make a prediction on the distribution of skill levels by firm or location. Any distribution would be consistent with equilibrium. Our first result is that in the absence of frictions and with social preferences, of however small strength, the equilibrium becomes skill

[1] For example, workers cannot post a bond, which would enforce the commitment to stay in the current firm.

segregated, that is, firms hire only from one skill pool.[2] The externality driving segregation is different than the one in models of say, racial segregation. We deal here with a pecuniary externality, that is, high-skilled types do not separate from low-skill types because they intrinsically dislike them. They do it, rather, because the market tends to produce different material payoffs for both.

Real markets are not perfectly frictionless, though. We introduce a particularly simple form of friction, moving (or hiring and training) costs, which produces additional implications on labor market outcomes. When moving costs between firms are low, heterogeneous productivities lead to widespread workplace skill segregation, and the whole market wage dispersion is explained by differences between firms. With intermediate levels of mobility costs, segregation is more moderate and wage dispersion arises both within and across firms. For high levels of moving costs the whole wage dispersion is within the firm, and becomes zero when the moving costs are sufficiently high. We show that within firm wage dispersion is associated with "internal labor market" features such as a dynamic form of wage compression, gradual promotions, and wage non-monotonicity.

These results arise from an interplay between risk preferences, social preferences and market competition. We examine these mechanisms separately.

We first discuss the implications in our model of the combination of risk preferences with our commitment structure. When there are *neither social preferences nor frictions*, the equilibrium labor contracts are as in Harris and Hölmström (1982), that is, wage payments are constant over time for a given observational type (for insurance reasons), and they change when the observational type changes. The presence of *frictions* in the market implies (in the absence of *social preferences*) that, when higher types are revealed, their wage changes less than in the absence of such frictions. Because of these frictions, workers remain employed with the firm that first hires them.

Next, we consider the effect of *social preferences and frictions*. As before, the frictions make it costly for workers to move between firms when their types are revealed. On the other hand, competitive pressure forces wages to be different for different (perceived) skill types. Thus, if workers of different types (who receive different wages) stay together, *social preferences* generate a loss in utility for some of them. To compensate for the disutility, the firm can increase the wages of the lower types.[3] The firm can also modify the composition of

[2]In a sense we can argue that social preferences operate here as a kind of "equilibrium-refinement." The advantage of this way of refining equilibria is that the payoff perturbation is economically and empirically well-motivated.

[3]Because of competitive pressures, there is no room to decrease wages for the higher types.

its workforce by letting some of the current workers leave and thus achieve a more homogenous (in terms of perceived skills) workers' pool. The firm now faces a trade-off between *wage compression* and *skill segregation*, and the size of the frictions determines the optimal solution to this trade-off.

The presence of frictions entices the market participants to find imaginative ways around them. In a sense, this is one of the lessons of the literature of contracting under incomplete information. One can exploit the diverse dimensions of preferences to "extract" private information by means of menus of contracts or nonlinear pricing.[4] This happens as well in our framework. Let us go back to the issue of the reference group for interpersonal comparisons. Recall that the reference group is not the whole set of firm employees, but only those that enjoyed similar circumstances in the near past. Then, *gradual promotions* appear as a less blunt tool than *wage compression* to lower the effects of social preferences. Rather than promoting an individual as soon as he is discovered to be of a high type, we show it is optimal to propose contracts which give a "smaller" promotion until his former peers "forget" him, and then promote him further later in the future. By doing so, the firm modifies the intertemporal composition of the reference group of each worker in a way that reduces the overall cost due to social preferences.[5] The dynamics of wages result from the complex interplay of the history of individual productivities, market competition and intra-firm reference group structure.

An additional implication of assuming "time-dependent" reference groups is that wage schedules may be *non-monotonic*. When some individuals' performances have started to differ only recently from others, there is some *wage compression*, raising the salaries of low types. Once the high types have disappeared from the reference group, the salary of the low types can fall back to "normal."

As one can readily see, social preferences produce a wide variety of effects that happen in well-specified circumstances, ranging from segregation by skill, to wage compression, gradual promotions and non-monotonic wages. Models with this richness allow for a better empirical fit with reality (if, as we expect, social preferences of this form are indeed present). They also suggest that labor and human resource economics can greatly benefit from incorporating behavioral factors in their standard set of tools.

[4]The standard example for competitive screening (see e.g. Mas-Colell, Green and Whinston 1995, p. 460) shows that firms use the differences in cost of effort (about which the firm does not care directly) to separate the workers of different productivity, and thus minimize the informational rents they extract.

[5]Because of insurance effects for the high type, this gradual promotion is second-best. We show that the firm balances this inefficiency with the social concerns to choose an optimal (gradual) promotion path.

Background and related work We bring together several strands of the economics literature.

Research on social preferences originated in large measure to give account of the growing empirical and experimental evidence that human behavior could not be explained only by the hypothesis of self-interested material payoff maximization. For instance, contribution to public goods is higher than would be expected under purely selfish maximization.[6] More importantly from our point of view, there is vast amounts of evidence that people reject lopsided offers in ultimatum bargaining games.[7] Several models have been proposed to account for these observations,[8] and we refer to the excellent surveys of Sobel (2000) and Fehr and Schmidt (2000b) for a discussion. A feature that many of the models share is that individuals dislike payoff inequality. One innovation with respect to this literature is that we think explicitly about the set of individuals to which the utility comparisons apply. In our paper, the reference group for comparisons is a product of the collective employment history. Workers identify less with superiors than with co-workers at their same level or recently promoted. Akerlof and Kranton (2000) also relate identity with incentive problems. In their case, the agents identification with a particular group gives them an incentive to exert effort, in a moral-hazard context. For us, the identification with a reference group creates disutility for individuals who earn less than the average in their reference group.

A few papers examine the implications for wages and the labor market of social preferences. Frank (1982) in his seminal paper showed that workers need not be paid their marginal productivity if people had preferences such that they cared sufficiently strongly (and in a heterogeneous way) about relative payoffs, liking to be better paid than others, and disliking to be paid worse. The more productive people would be paid less than their marginal productivity as they got the "pleasure" of earning more than their colleagues. Similarly, the less productive people would be paid more than their marginal productivity so as to be compensated for the "suffering" of earning an inferior wage.[9] In our paper we add a dynamic

[6]See Ledyard's (1995) survey on public goods in the *Handbook of Experimental Economics*.

[7]See Güth, Schmittberger and Schwarze (1982) and also Roth's (1995) survey on bargaining in the *Handbook of Experimental Economics*.

[8]Bolton (1991), Rabin (1993), Levine (1998), Bolton and Ockenfels (1999), Fehr and Schmidt (2000a), Charness and Rabin (2002).

[9]Frank (1985) discusses many practical implications of this basic framework, such as the puzzling omnipresence of minimum wages, safety regulations, forced savings for retirement and other labor market regulations. These can be explained with his model as a way to compensate for the externality that is generated by the social preferences. Other papers which deal with contracting problems and social preferences are Fehr, Klein and Schmidt (2001), Fershtman, Hvide and Weiss (2003) and Rey-Biel (2002).

dimension to the contracting problem. This allows us to discuss issues such as promotions, worker flows and the evolution of wages, which widens the span of testable implications. Besides, this is probably the first model to characterize long-term contracts in the presence of social preferences in a competitive labor market.

There is also evidence that firms workforces are more homogenous than simple "random matching" would suggest. People of different skill levels sort themselves into different firms. For instance, Kramarz, Lollivier and Pelé (1996) find that specialization[10] increased massively in France between 1986 and 1992.[11] Davis and Haltinwanger (1991) note that the continuous rise in wage inequality in the U.S. is imputable in part to ability sorting of workers across firms. Brown and Medoff (1991) investigate explanations for wage-size differentials, and find evidence in support only for explanations based on sorting by worker skill. Theoretical explanations for this evidence usually resort to the introduction of some form of complementarities between individuals of the same skill levels.[12] We depart from this by not postulating any form of production complementarities between worker's types. The externality that arises between workers is of a pecuniary nature. It arises because market outcomes favor more productive workers, and individuals are averse to inequalities in their own reference group.[13] Besides, our model suggests that the time-series evidence on skill segregation can be related to changes in labor market regulations (and organizational features) that affect mobility costs.

Bewley (1999) offers direct evidence for the kind of externality we postulate. Some 78% of the businesspeople whom he asked about internal equity, say that it is important for internal harmony and morale.[14] Morale meant "cooperativeness, happiness or tolerance of unpleasantness, and zest for the job."[15] Section 6.5 in Bewley (1999) has a number of

[10]They compute a measure of specialization for different professional categories as proposed by Kremer and Maskin (1996).

[11]"Blue collar unskilled workers are more and more separated from other types of workers, and therefore, tend to work together in the same firms. This is true for each of the six categories of skills. The number even doubled for clerks." Kramarz *et al.* (1996), p. 375.

[12]Good examples of these explanations are de Bartolomé (1990), Bénabou (1993), Kremer and Maskin (1996) and Saint-Paul (2001). The theoretical papers of Legros and Newman (2002, 2004) identify the minimal conditions for such positive sorting.

[13]There are other models of segregation which rely on group externalities. Seminal works in this area are Becker (1957) and Schelling (1971). Contrary to our paper, in that literature the individuals have an intrinsic like or dislike of workers in their or other groups. In our case, the spillover is related only to the market outcome. High and low types would live happily together if wages were equal.

[14]Bewley (1999), table 6.5.

[15]Bewley (1999) p. 42.

revealing quotes from managers about the disruptive effects of lack of equity on the job.[16] He also shows that an important consequence of internal inequity in firms is turnover,[17] just as our model would predict.

We also predict that social preferences lead to *wage compression*. By this, we mean the differences in wages between workers of different perceived skills is lower with than without social preferences.[18] The evidence given in the literature for wage compression is often indirect. The ratio between, say, the lowest 10th percentile and the highest 90th percentile of the wage distribution has undergone dramatic variation over time, and is quite different between countries, in a way that is hard to justify from purely technological reasons.[19] The study of Cannon, Fallick, Lettau and Saks (2001), which directly compares wages and productivity, shows that wage compression may arise, as in our model, "from the value workers place on relative pay," (p. 3). On the other hand, Hibbs and Locking (2000) show that within plant and within industry wage leveling adversely affected productive efficiency in Sweden. This contradicts an explanation of *wage compression* as the result of firms trying to enhance the morale of workers, thus achieving higher productivity, as in Akerlof and Yellen (1990), but it is still consistent with our approach.

Section 2 describes the dynamic labor market model. Section 3 presents the recursive formulation of the problem and states the equivalence with the market game in Section 2. All results are gathered in Section 4. Appendix A describes the recursive formulation in its most general form, and establishes the equivalence between the market game in Section 2 and the simplified recursive formulation in Section 3. The proofs of the results stated in Section 4 are in Appendix B. Appendix C contains an exhaustive analysis of the value function of our model.

2 Model

Time is discrete and indexed by $t = 1, ..., T < \infty$.

[16] ¿From "Internal equity is very important," to "Inequity causes disharmony" and even "Unfairness can cause upheaval within an organization and lead to disfunctional activities."

[17] Bewley (1999), table 6.5.

[18] Another model predicting *wage compression* in the presence of mobility costs is Acemoglu and Pischke (1999). They show that wage compression is a necessary condition for firms' investment in general training. Acemoglu (1999) relates *wage compression* to search frictions.

[19] Classic references in this context are Katz and Murphy (1992) or Goldin and Margo (1992).

Firms There is a finite set M_t of risk-neutral firms who enter the market and post an offer at any period. For simplicity, and w.l.o.g.,[20] we assume that the M_ts are disjoint, so that $M^t = \cup_{s=1}^{t} M_s$ is the set of firms that had the chance to make an offer at some date prior to t.

Firms that are active in the market collect profits at the end of each period. Firms discount at zero interest rate.

Workers and timing Workers are risk averse and live for T periods.

There is a continuum of workers in $[0, 1]$ of two different types, g(ood) workers and b(ad) workers. Workers g produce one unit of output per period with *i.i.d.* probability p (and zero otherwise), while workers of type b have no chance of producing good outcomes (their production is always zero). We denote by $\lambda \in (0, 1)$ the number of workers of type g in the population. Information about workers' types is imperfect but symmetric, as in Harris and Hölmström (1982).

In each period t the timing of payment is as follows. The worker decides wether to stay in the firm or accept an outside offer. If the worker decides to stay in the firm, he receives the wage from his employer.[21] He then produces (thereby possibly revealing his type). This new information is then used at the beginning of the next period by the entrant firms (the market) to make job offers and by the old firm to pay $t + 1$ wages taking into account the labor market pressure.

Worker assignment At each period t, the mapping $f_t : [0, 1] \to M^t \cup \{0\}$ keeps track of the assignment of workers to firms. The case $f_t(i) = 0$ corresponds to worker i being unemployed.

Whenever a worker changes firm he pays a fixed mobility cost $k \geq 0$. This can be interpreted as a moving or hiring cost.

Outputs and types Firms learn about the workers' types by observing production outcomes of each period.

Consider some worker $i \in [0, 1]$. Let $y_t^i = 1$ if worker i generates a positive output at t (thus revealing he is of type g), and $y_t^i = 0$, otherwise. We set $y_0^i = 0$.

The quality of the worker is a crucial state variable of this problem. Let q_t^i be the belief that a worker i is of good type at beginning of period t. By Bayes' rule, next period's quality

[20] Indeed, we assume that firms' offers commit them for future periods.

[21] We will see that, for insurance purposes, the worker may, in fact, receive (severance) payments from earlier employers as well.

value is $q^i_{t+1} = 1$ if $y^i_t = 1$ and

$$q^i_{t+1} = \frac{q^i_t (1-p)}{q^i_t (1-p) + (1-q^i_t)}. \tag{1}$$

if $y^i_t = 0$, with initial condition $q^i_1 = \lambda$ for all i.[22]

Contracts and contract offers A long term contract specifies a sequence of non-negative payments contingent on observed history which includes worker-firm assignments, production and types. There is full commitment from the part of the firm on the terms of the contract.

Let $h^i_t = \left\{\left(y^i_{s-1},\ q^i_s,\ f_s(i)\right)\right\}^t_{s=1}$ be the agent i's individual history at the beginning of period t, after period t employer has been chosen. Let H be the set of all conceivable histories. For all $s \geq t$, denote by $H(h_t; s)$ the set of histories starting from node h_t (including node h_t) until period s. In our model, individuals are identified with their histories. Let $(h_t \backslash f_t) = (h_{t-1}, y_{t-1}, q_t)$ be a shorter notation for history h_t without specifying the worker-firm assignment at period t.

Definition 1 *A feasible contract offer $W^j_t(h_t \backslash f_t)$ in period t by firm $j \in M_t$ is a collection $\{w^j_s\}^T_{s=t}$ of mappings $w^j_s : H(h_t; s) \times [0,1] \to I\!R_+$ such that, for all $i \in [0,1]$ and $h_s \in H(h_t; s)$, $w^j_{s,i}(h_s)$ is the wage paid in period $s \geq t$ to worker i. Moreover, we assume that if $h^i_t = h^{i'}_t$, then $w^j_{s,i}(\cdot) = w^j_{s,i'}(\cdot)$, for all $s \geq t$.*

Notice that we assume that firms cannot post contracts offers that depend on the identity of the worker *per se*, but we allow them to depend on each past worker's employment history. From now one, and to simplify notation, we thus omit the worker index in the payment schedules.

At each period, all firms simultaneously post feasible contracts, taking as given previous offers.

Then, workers simultaneously decide whether to accept any new contract, to remain with the current employer at the previously agreed contract, or to go unemployed.

At each period t there will be old contract in place as well. When $j \in M_s$ for some $s < t$, then $W^j_t(h_t \backslash f_t)$ simply denotes the continuation of the contract W^j_s after node h_t. Unemployment corresponds to an offer that pays zero under any present or future contingency. Let $\mathcal{W} = \left\{W^{j_t}_t(\cdot)\right\}^{j_t \in M^t \cup \{0\}}_{t=1,\ldots T}$ be the whole set of contract offers.

[22]This stochastic structure of types implies that $q = 1$ is an absorbing state. This way we simplify the nature of contracts but the intuition carries over with a richer stochastic structure.

Workers' strategies Worker-firm assignments are determined by workers' decisions in any period. We represent the choice of worker i by a sequence of functions $F_t^i(\cdot)$ of the form $F_t^i : H(h_t \backslash f_t) \rightarrow M^t \cup \{0\}$. Denote by $\mathbf{F}^i = \{F_t^i(\cdot)\}_{t=1}^T$ a complete sequence of such functions, which completely describes worker i's choice.

These functions are essential in generating individual histories. Consider such a sequence. Then, according to this sequence, at period 1, worker i goes to firm $f_1(i) = F_1^i(0, \lambda)$ and the resulting individual history is $h_1^i = \{\emptyset\} \cup (0, \lambda, f_1(i))$.[23] Then, production y_1^i takes place, and firms update their beliefs to q_2^i at the beginning of period 2. The firm assignment $f_2(i)$ of worker i in period 2 is determined by the mapping $F_2^i(h_2^i \backslash f_2) = F_2^i(h_1^i, y_1^i, q_2^i)$, and so on until period T. Notice that since the unemployment offer is always in place, each worker's decision function $F_t^i(\cdot)$ is well defined at each node.

Denote by $\mathcal{F} = \{\mathbf{F}^i(\cdot)\}_{i \in [0,1]}$ the whole set of workers' assignment sequences.

Workers' (social) preferences Notice that the set of contract offers $\mathcal{W}$ generate "total wage" schedules $w_t(\cdot) : H \rightarrow \mathbb{R}$ defined a follows:

$$w_t(h_t) = \sum_{j \in M^t} w_t^j(h_t). \tag{2}$$

Let $\mathbf{w} = \{w_t(\cdot)\}_{t=1}^T$ be a set of "total wage" functions.

In addition to the utility they obtain from their own wage $-$their material payoffs$-$ workers also experience (dis)utility from the material payoffs of firm mates in their reference group. More precisely, if we let $w_t(h_t)$ be the worker wage at node h_t, his instantaneous utility at period t is:

$$u(w(h_t)) - A(\overline{w}_t(h_t) - w_t(h_t)),$$

where $\overline{w}_t(h_t)$ is the maximum of h_t's firm mates' wages in his reference group and $A(\cdot)$ is the function expressing the aversion to inequity.

We assume that $A(\cdot)$ is zero-valued for $x \leq 0$, non-decreasing for $x > 0$, continuously differentiable with $A'(0) = 0$, and convex. For instance, $A(x) = \alpha \max\{x, 0\}^p$, with $p > 1$ and $\alpha \geq 0$. Under these conditions, i experiences a disutility if and only if h_t's co-workers highest wage is higher than his own.[24] The material payoff is described by a strictly concave and differentiable utility u.

[23] Recall that, by assumption, $h_0^i = \emptyset$, $y_0^i = 0$,and $q_1^i = \lambda$ for all worker i.

[24] Technically, this is an extreme version of difference aversion models such as Bolton and Ockenfels (2000) and Fehr and Schmidt (2000a). There are other models of social preferences were agents care about the actions of others (reciprocity). See, for example, Levine (1998) and Charness and Rabin (2002).

We assume that the reference group of worker h_t at period t corresponds to the set of h_t's co-workers of same type than h_t at period $t-1$. In other words, the reference group is not the whole set of firm employees, but only those that enjoyed similar circumstances in the near past.[25]

Denote by $R_t(h_t)$ the reference group of worker h_t at t. We have

$$R_t(h_t) = \left\{ h_t' \in f_{t-1}^{-1}(h_t) : \; q_{t-1}' = q_{t-1} \text{ and } f_t' = f_t \right\}. \tag{3}$$

where q_{t-1} is the quality of worker h_t in period $t-1$ and f_t and f_t' are the period t employer of worker h_t and h_t' respectively (that is, the last entries in h_t and h_t'). Then, if $R_t(h_t)$ has positive mass we have

$$\overline{w}_t(h_t) \equiv \sup_{h_t' \in R_t(h_t)} w_t(h_t') \tag{4}$$

which defines a max wage schedule relevant for social preferences. We assume that $\overline{w}_t(h_t) = w_t(h_t)$ each time $R_t(h_t)$ has zero mass (including, obviously, the case when $R_t(h_t) = h_t$).

In equilibrium, rational agents compute the max wage function $\overline{w}_t(\cdot)$ using contract offers $\mathcal{W}$ and allocation rules $\mathcal{F}$. Let $\bar{\mathbf{w}} = \{\overline{w}_t(h_t)\}_{t=1}^T$ be the set of such max wage functions.

Given a $\mathbf{w}$ and a $\bar{\mathbf{w}}$, by choosing a set of assignment decision rules $\mathbf{F}$ workers assign a lifetime utility value to each node $h_t \backslash f_t$ in the usual way:

$$U_t(h_t \backslash f_t, \mathbf{F}; \mathbf{w}, \bar{\mathbf{w}}) = \mathbf{E}\left[\sum_{n=0}^{T-t} u(w_{t+n}(h_{t+n})) - A(\bar{w}_t(h_{t+n}) - w_t(h_{t+n})) \mid h_t \right]. \tag{5}$$

Notice that the expectation operator is always well defined since $\mathbf{F}$ specifies history h_t which follows node $h_t \backslash f_t$, even for nodes which are non consistent with $\mathbf{F}$. When the other arguments are unambiguously defined we will denote by $U_t(h_t)$ a function which associates lifetime a utility value to each node $h_t \in H$. Let $\mathbf{U} = \{U_t(\cdot)\}_{t=1}^T$ be a set of such functions.

Definition 2 (Equilibrium) *An equilibrium outcome is a tuple* $[\mathcal{W}, \mathcal{F}, \mathbf{U}, \mathbf{w}, \bar{\mathbf{w}}]$ *with the following properties:*

(*i*) *Profit maximization:* $\mathcal{W}$ *is such that, given the assignment* $\mathcal{F}$, *and* $\mathcal{W} \backslash \{W_t^j(h_t \backslash f_t)\}$, *each new firm* $j \in M_t$ *maximizes its expected profits at* $W_t^j(h_t \backslash f_t)$;

(*ii*) *Optimal assignment:* $\mathcal{F}$ *is such that each worker* i *maximizes his lifetime utility (5) at* $\mathbf{F}^i$ *taking as given* $\mathbf{w}$ *and* $\bar{\mathbf{w}}$;

[25] More generally, we could have assumed that the reference group of worker i at period t is equal to the set of i's co-workers of same type than i at periods $t-1$ to $t-r$ (for some fixed r), with possibly different weight for each group. This extension would considerably enlarge the state space, but all our results will hold with this more general specification as well.

(*iii*) *Rational Expectations:* $\mathbf{w}$ *and* $\bar{\mathbf{w}}$ *are computed from* $\mathcal{W}$, $\mathcal{F}$ *using (2), and (3) and (4), respectively.*

The optimal assignment strategies $\mathcal{F}$ can be constructed backward as follows. Recall that $f_{T-1}(i)$ is the firm that employed worker i at period $T-1$. Let $h_T \backslash f_T$ be a last period node before firms make offers. At each such node, the worker decides to remain inside the firm or to leave by joining a competitor. Formally, worker i solves:

$$\max_{\rho \in \{0,1\}} \rho U_T(h_{T-1}, q_T, f_{T-1}) + (1-\rho) U_T^m(h_T \backslash f_T).$$

where $\rho = 1$ (resp. $\rho = 0$) stands for staying in (resp. leaving) the current firm. The expression $U_T^m(h_T \backslash f_T)$ corresponds to the best market offer, that is:

$$U_T^m(h_T \backslash f_T) = \sup_{j \neq f_{T-1}(i), j \in M^t \cup \{0\}} U_1(h_{T-1},\ q_T,\ j).$$

At equilibrium, $f_T(i) = f_{T-1}(i)$ if and only if $\rho^* = 1$. Otherwise, the identity $f_T(i)$ of the new employer coincides with any of the best market offers available.[26]

3 Optimal contracts: a recursive formulation

In this section we show that the equilibrium allocation can be characterized recursively. This is so because firms' full commitment and the possibility of paying severance payments make the equilibrium constrained efficient.

For this purpose, we begin by formulating a recursive constrained optimization problem. First, we define the optimization problem for the last period. This is of course required because our setting has a finite time horizon. Besides, the last period formulation allows us to discuss with detail the constraints of the problem. Then, we present the general recursive expression. Finally, we show that the equilibrium of our game coincides with the solution of this optimization problem.

We first introduce some useful notations.

Good workers Consider some worker i that is known to be good at the beginning of time t (that is, $q_t^i = 1$). Then, from period $t+1$ on, the reference group of worker i includes only workers' of good type. That is, for $s \geq t+1$, only workers j with $q_s^j = 1$ are in his reference group, and he is only in the reference group of such workers. Because of this and given our definitions of social preferences, there are no externalities across these workers and

[26] Ties are broken randomly.

any other any more. Thus standard arguments imply that market competition and worker's risk aversion produce for all $s > t$ equilibrium wages for these workers that are equal across periods and production realizations (for insurance reasons), and across workers.

Workers of yet unknown type Consider a firm f designing the contingent payments to be effective at the end of period t. Workers that were not working in f in the previous period do not belong to the reference group for workers already in f at $t-1$ and vice versa. Thus, their contracts can be treated separately from the point of view of firm f. Similarly, as we argued before, workers i with $q^i_{t-1} = 1$ can be treated separately as well.

Thus we only need to focus on the characteristics of contracts for workers with $q^i_{t-1} \neq 1$, and who were employed in firm f at $t-1$. We denote by w_{gt} (respectively, w_{ut}) the wage of such a worker when $y^i_t = 1$ (respectively, $y^i_t = 0$). It follows from our definition that these payments are independent of the identity i of the worker and that of his employer f.

The recursive formulation: last period Whenever no confusion is possible, and to simplify notations, we use letters without time subscripts to denote choice variables.

We consider first the firms's problem in the last period (at date $T-1$). Let q be the average quality of the current pool of workers *within the same reference group* (agents who had the same past history till the last period and enrolled in the same firm). Denote by π the profits the firm makes out of this group. If we denote by $V_T(\pi, q)$ the *ex-ante* (before production of the previous period realizes) utility of a worker who belongs to this poll. We have:

$$V_T(\pi, q) = \max_{w_u, w_g, \rho} \; pqu(w_g) + (1-qp)\left[u(w_u) - \rho A(\widetilde{w}(q))\right]$$

subject to

$$\begin{aligned} \rho \in \{0,1\} \qquad & [\rho] \\ q_u = \tfrac{q(1-p)}{q(1-p)+(1-q)} \qquad & [q] \\ \widetilde{w}(q) = w_g - w_u \qquad & [\alpha] \\ u(w_u) - \rho A(\widetilde{w}(q)) \geq u(pq_u - k) \qquad & [u] \\ u(w_g) \geq u(p-k) \qquad & [g] \\ pq\,(p - w_g) + (1-pq)\,(pq_u - w_u) - (1-\rho)\min\{pq, 1-pq\}k \geq \pi \qquad & [\pi] \end{aligned}$$

We comment on this optimization problem.

Equation $[q]$ is the Bayes' rule (1) for the average quality of the workers of still unknown type.

Equation $[\alpha]$ computes the difference between the wage w_u of workers of unknown type in the firm and the workers of good type. This difference is the source of the social preferences disutility (thus, cost for the firm).

Equations [u] and [g] are the participation constraints of, respectively, workers of type u and g. The left-hand side is simply the utility of accepting the proposed contract. The right-hand side is the utility derived from the market wage. This market wage results from the zero-profit condition for the highest bidding entrant. More precisely, we know that workers of type u generate on average pq_u this period. The zero profit condition for these workers is then:

$$pq_u - w_u^* = 0,$$

where w_u^* is the market wage. However, since these workers must pay the moving cost k we can assume that this cost is transferred to the firm. Thus, firms' profits are equal to k and workers' net earnings are $w_u^* - k = pq_u - k$. Similarly, the market wage of type g workers is such that $p - w_g^* = 0$, and net earnings are $w_g^* - k = p - k$. Notice that since $q_u < 1$ the right hand side of [g] will always be larger than the right hand side of [u]. This is why good workers will never be affected by social preferences.

We now explain the constraint [ρ]. In our context, firms offer long-term contracts and face competition by entrants. Thus, a firm who keeps workers of the same type together faces a cost due to workers' social concerns. On the other hand, firms are somewhat shielded from competition (thus pay slightly lower wages) because of the moving costs that a competitor needs to pay in order to steal new workers. Hence, keeping workers in the firm is a matter of choice, and the probability of keeping the worker ρ models this choice. The main trade-off here is between the cost (higher wages) generated by social concerns and the benefit (lower wages) arising from the hiring cost. In principle, a firm might want to let go either the good types, or the unknown types, or both, depending on the circumstances.[27] We show in the appendix that the firm always fires the workers from the smaller-sized pool (either the good or those of yet unknown type), and thus one can formulate the problem with only one ρ, that keeps track of whether somebody is fired at all.

Equation [π] guarantees that, with the proposed wage contract, the firm can secure expected profits at least equal to π.

The recursive formulation: the general case Let $V_t(\pi, q)$ be the *ex-ante* utility of a worker who belong to a reference group of average quality q, when the employer is expecting to make an ex-ante level of profits equal to π when there are $T - t \geq 0$ periods before the

[27] For example, the cost due to social concerns varies with workforce composition. By taking different decisions as to which type of workers leave, one can modify the workforce composition and, thus, change this cost.

end. Obviously, $V_T \equiv 0$, in general we have:

$$V_t(\pi, q) = \max_{w_u, w_g, \rho, \pi_u, \pi_g} \left\{ \begin{array}{c} (1-pq)\left[u(w_u) - \rho A(\widetilde{w}(q)) + V_{t+1}(\pi_u, q_u)\right] \\ +pq\left[u(w_g) + V_{t+1}(\pi_g, 1)\right] \end{array} \right\}$$

subject to

$$\begin{array}{rr} \rho \in \{0,1\} & [\rho] \\ q_u = \frac{q(1-p)}{q(1-p)+(1-q)} & [q] \\ \widetilde{w}(q) = w_g - w_u & [\alpha] \\ u(w_u) - \rho A(\widetilde{w}(q)) + V_{t+1}(\pi_u, q_u) \geq V_t^m(k, q_u) & [u] \\ u(w_g) + V_{t+1}(\pi_g, 1) \geq V_t^m(k, 1) & [g] \\ pq(p - w_g + \pi_g) + (1-pq)\left(pq_u - w_u + \pi_u\right)) - (1-\rho)\min\{pq, 1-pq\}\, k \geq \pi, & [\pi] \end{array}$$

where

$$V_t^m(k, q) = \max_{w, \pi} \left\{ \begin{array}{cc} & u(w) + V_{t+1}(\pi, q) \\ s.t. & pq - w + \pi - k \geq 0 \end{array} \right\}$$

is the maximal utility obtainable in the market by a pool of workers of quality q. Hence, for good workers ($q = 1$) we have:

$$V_t^m(k, 1) = \max_{w, \pi} \left\{ \begin{array}{cc} & u(w) + V_{t+1}(\pi, 1) \\ s.t. & pq - w + \pi - k \geq 0 \end{array} \right\}$$

It is easy to see that the problem for good workers is fully stationary. In this case the market contract consists of a constant wage $w_{t,g}^* = p - \frac{k}{T+1-t}$. Indeed, when $q = 1$, there is no further heterogeneity in the pool, and hence $V_t(\pi, 1) = V_t^m(\pi, 1)$ for all $\pi,\ t$. Then, $V_t(\pi, 1)$ is a strictly concave and differentiable function, with:[28]

$$\frac{\partial}{\partial \pi} V_t(\pi, 1) = u'\left(p - \frac{\pi}{T+1-t}\right),$$

which leads to the expression for the wage.

The equivalence result The following result guarantees that we can solve for an equilibrium of the game in Section 2 by characterizing the solution to the optimization problem defined above. Besides, existence of the equilibrium is guaranteed.

Proposition 1 (recursive equivalence) *An equilibrium of the game described in Section 2 always exists. Let the policy functions* $\mathcal{O}(0, \lambda) = \{w_t, \pi_t, \pi_{z,t}, w_{z,t}, \rho_t\}_{z=g,u;t=1,..,T}$ *and the*

[28] See Appendix C.

value functions $\mathcal{V}(0,\lambda) = \{V_t, V_t^m\}_{t=1,..,T}$ *be a solution to the maximization problem described in Section 3 when* $\pi_1 = 0$ *and* $q_1 = \lambda$. *Then, in any undominated equilibrium* $[\mathcal{W}, \mathcal{F}, \mathbf{U}, \mathbf{w},\ \bar{\mathbf{w}}]$ *of the game the wage offers, firm-worker assignments and payoffs of all workers (except for at most a measure zero set of them) are given by* $\mathcal{O}(0,\lambda)$ *and* $\mathcal{V}(0,\lambda)$.

At this equilibrium, the *ex-ante* utility of a worker belonging to a reference group of average quality q is $\mathbf{E}\left[U_t\left(h_t\right) \mid h_{t-1}\right] = V_t(\pi, q)$.

4 Optimal contracts: the results

4.1 The case without social preferences

We aim to understand the effect of social preferences on the allocation of workers to firms and on the wage profiles. For this purpose we first describe the predictions of our model in the absence of social concerns, that is, when $A \equiv 0$. In this case, the model extends Harris and Hölmström (1982) (HH hereafter) to a setting with mobility costs, $k \geq 0$. The case when $k = 0$ is a discrete support of human capital levels version of HH.

Proposition 2 (no social preferences) *Assume that there are no social concerns (*$A \equiv 0$*). Then:*

(*i*) *when* $k = 0$*, the firm-worker assignment is indeterminate;*

(*ii*) *when* $k > 0$*, no worker ever leaves his initial employer, that is,* $\rho_t = 1$ *for all* t*;*

(*iii*) *for all* $k \geq 0$*, the wage schedule is downward rigid, that is,* $w_{z',t+1} \geq w_{z,t}$ *for all* t *and* $z, z' \in \{g, u\}$*;*

(*iv*) *wages are stationary for a given type, that is,* $w_{z,t+1} = w_{z,t}$ *for all* t *and* $z \in \{g, u\}$.

For very large levels of k, the market pressure is so low that the firm can fully insure the workers and pay them their expected productivity each period, i.e. $w_t\left(h_t\right) = \lambda p$ for each equilibrium history h_t. For more moderate levels of moving costs (including $k = 0$) the model generates monotone (downward rigid) wages. In period 1 each worker is paid less than his expected productivity and the wage remains constant until the worker is revealed to be good. When the worker's type is revealed, he will be approached by an external firm, and his wage within the original firm must increase to match the market offer. His wage remains constant from that period onwards. Notice that when $k > 0$, there are neither quits nor layoffs, and when $k = 0$, worker flows are indeterminate.

4.2 The case with social preferences and without mobility costs

¿From now on we will consider the case that there are social preferences, that is, $A(x)$ strictly increasing when $x > 0$. When $k = 0$, we have a full segregation result, i.e. there will be no workers' heterogeneity within the same firm, and all wage dispersion is between firms.

Proposition 3 (skill segregation) *If $k = 0$, then $\rho_t = 0$, for all t, that is, firms hire from only one skill pool. Hence wage dispersion within the firm is zero, but overall wage dispersion is maximal and identical to the case without social preferences described in Proposition 2 (with $k = 0$).*

In the absence of mobility costs, segregating the workforce saves on the pecuniary externality created by competitive pressures and the presence of social concerns within firms. In other models which produce segregation, this is driven by a direct externality over others' attributes.[29] Agents, say, have preferences over the types of others. Here, preferences are only indirectly affected by the types of others, as the primary externality is induced by economic outcomes (which are, in turn, shaped by differences in type productivity and competitive pressures).

A corollary of this result is that worker compensation in this framework has the same structure as in HH. The good type, which has completely revealed his type, receives his expected productivity. For the other type, compensation are downward rigid and trade off the insurance concern of the risk-averse agents with the competitive pressure. Insurance creates a tendency to have constant wages. But since workers are free to move between firms, the good types necessarily have to be compensated when they reveal their type. The key difference with respect to HH is that here some workers actually leave the firm, i.e. this model produces worker flows. In addition, notice that social preferences imply that in this extreme case all the observed wage dispersion is between firms. Within the same firm all workers receive the same wage. Finally notice that if those who leave are not the good types, they might be entitled to a compensation that is higher than their expected productivity. Since the new firm does not pay a wage in excess of expected productivity, the difference is made up by the former employer, in the form of severance payments.[30]

Notice that also in the case with social preferences for very large levels of k, the market pressure is so low that the firm can fully insure the workers and pay them their expected

[29] Seminal works in this area are Becker (1957) and Schelling (1971).

[30] The presence of severance payments allows the firm to pay smaller wages during the employment period. In this way, one can reinterpret this payment (and the lower wages in the past) as an optimal unemployment insurance scheme (see Hopenhayn and Nicolini 1997 and Pavoni 2004).

productivity each period, i.e. $w_t(h_t) \equiv \lambda p$.

4.3 The general case

We now consider the general case with both social concerns and mobility costs. That is we assume that $A(x)$ is strictly increasing when $x > 0$, and that $k > 0$. With respect to the previous case (where $k = 0$), the introduction of frictions in the form of mobility costs creates a trade-off between the gains in efficiency stemming from lower inequality when workers are free to move between firms and the loss created by the hiring cost associated with these moves. The wage structure will subsume this source of inefficiencies in three different ways, which we analyze in turn.

First, through lower intra-firm wage dispersion. When the cost of mobility is not too small there is heterogeneity in the perceived skills within firms, which induces differential wages because of competitive pressures. Because of *social preferences* workers of yet unknown type suffer a loss in utility. To compensate them (thus avoiding quits to the competitors), the firm pays them a higher wage than that they would receive in the absence of social concerns.[31] We call this effect, *wage compression.* Thus, social concerns add a new source of wage compression in addition to the one already derived from insurance.

More formally,

Proposition 4 (wage compression) *Assume that at two successive dates with $\rho_{t+1} = 1$, we have $0 \leq \widetilde{w}(q_t) < \widetilde{w}(q_{t+1})$, then $w_{u,t} < w_{u,t+1}$. In particular, when social concerns are not active in the current period but are active next period, the next period wage for workers of yet unknown type is larger than their wage in the current period.*

The second observational implication of frictions in our model is that the reaction (in term of wage increases) to new positive information about workers will be more gradual than one would expect from pure market forces. We have assumed that the reference group within which social concerns are active is composed by co-workers who recently were in similar circumstances. An immediate reaction of wages to productivity followed by a flat wage scheme is optimal from the point of view of intertemporal smoothing of the worker whose good type has just been revealed. This one-step wage increase is, on the other hand, costly from the point of view of social concerns. It would be preferable to do a more gradual increase, taking advantage of the fact that any wage increase, however small, would separate

[31] The higher wage for the unknown type reduces the wage differential that creates social concern. In particular w_u increases even when the constraint $[u]$ is slack.

the lucky worker from the reference group of the less fortunate ones. This creates the scope for reducing the cost of inequality by making the transitions more gradual.

We call this *gradual promotions*. This is a qualitatively new feature of the wage dynamics, where firms exploit an endogenous dimension of the worker's preferences, the reference group, which they manipulate through the reaction of wage patterns to output realizations.

Notice that −as Proposition 2*(iv)* shows− this feature of the wage profile is generated in our model by social preferences, i.e. it is not present in the model when $A \equiv 0$.

Proposition 5 (gradual promotions) *If $\widetilde{w}(q_t) > 0$, and $\rho_t = 1$ then $w_{g,t+1} > w_{g,t}$, for a worker revealed to be good at time t (that is, for worker i such that $y_s^i = 0$, for all $s < t$, and $y_t^i = 1$). In other words, when the wage of the workers of the good type is meant to increase because the participation constraint is binding and social concerns are active in the current period, the wage for workers of the good type increases gradually towards their (known) productivity.*

A final observation regarding the dynamic pattern of wages is that they do not need to be monotone, unlike in HH, where wages are downward rigid. We find that wages can decrease after an expansive phase, because the reference group for wage comparison changes during this phase, and so do the social concerns that condition the wages that are paid. We call this *wage non-monotonicity*.

Proposition 6 (wage non-monotonicity) *Consider three successive dates with $\rho_t = 1$, and both $\widetilde{w}(q_{t-1}) < \widetilde{w}(q_t)$ and $\widetilde{w}(q_{t+1}) < \widetilde{w}(q_t)$. Then, if constraint [u] is not binding at period $t+1$, both $w_{u,t-1} < w_{u,t}$ and $w_{u,t+1} < w_{u,t}$. In particular, if social concerns are not active in periods $t-1$ and $t+1$ but are active in period t, the wage increases between $t-1$ and t and then decreases in period $t+1$.*

Notice again that this result is not present in the model with no social concerns, and that this is a characteristic of the first three periods of a relationship.[32]

5 Final remarks

This paper provides a new dynamic competitive equilibrium model of the labor market. The presence of social concerns and mobility costs has both cross section and time series

[32]After the third period, the wage of the unknown type decreases as long as $k > 0$ because of the finite time-horizon effect.

implications for the market allocations of workers to firms, for within and between firms wage dispersion, and for the internal wage structure of the firm.

First, social preferences generate workplace skill segregation, whose extent decreases with mobility costs. This prediction is consistent with the widening of inter-firm wage variance observed in the last decade (see Kramarz, Lollivier, and Pelé 1996 for France and Kremer and Maskin 1996 for the U.S.). Second, social preferences and mobility costs reduce within firms' wage variance (documented empirically by Goldin and Margo 1992, or Katz and Murphy 1992). Third, individual wage changes at the firm level are serially correlated, consistently with the findings of Baker, Gibbs and Hölmström (1994). Fourth, unlike in Harris and Hölmström (1982), wages can decrease within the firm in our model. This decreases are correlated with reorganizations and/or absence of promotions.

The model, thus, generates a broad range of implications. Lessons drawn initially from the experimental laboratory, once incorporated into standard models of organizations and markets, provide new quantitative and qualitative predictions which enrich our view of how the labor market operates.

Appendix A: Proof of Proposition 1.

The aim of this section is to establish the equivalence between the original equilibrium problem with social preferences and workers assignment decisions (the game of Section 2) and the recursive formulation in Section 3. We will then prove some properties of the associated value function in Appendix C.

Proof of Proposition 1 (recursive equivalence):

The existence of an equilibrium $[\mathcal{W}, \mathcal{F}, \mathbf{U}, \mathbf{w}, \bar{\mathbf{w}}]$ is established constructively.

First, at period $t = 1$, undominated contract offers must be such that payments from firm j are zero for each history h_1 such that $j \notin f_1([0,1])$. Hence, we can assume that only one firm makes non-zero payments at each h_1. Then, standard arguments imply that $U_1(h_1) = V_1^m(0, \lambda)$, that is, worker payoffs at the beginning of the game correspond to the market threat. Indeed, if a contract offer $W_1(0, \lambda)$ accepted in equilibrium by some worker did not solve the following problem:

$$V_1^m(0, \lambda) = \max_{w,\pi} \left\{ \begin{array}{cc} & u(w) + V_2(\pi, \lambda) \\ s.t. & p\lambda - w + \pi \geq 0 \end{array} \right\}, \tag{6}$$

then, either the offering firm would not make non-negative profits (the budget constraint above would not be satisfied)[33], or the offer would not maximize the worker's utility, or both. In all cases this would generate a contradiction for the following reasons. First, the fact that an equilibrium offer cannot generate negative profits ex-ante is immediate from the definition of equilibrium.[34] Second, if the offer W were not utility maximizing, there would exist another offer with positive profits delivering a higher utility to all workers; some competitor would make this offer and attract these workers.

Consider now a history h_t and consider the set of workers i for which $q_t^i \neq 1$. We denote by $\rho_{u,t}$ (resp. $\rho_{g,t}$) the symmetric equilibrium decision of all workers i such that $y_t^i = 0$ (resp.

[33]Notice that π reppresents the expected profits of the firm regardless of the details of the future offers.

[34]Profits can be computed as follows. Given the specified equilibrium, let

$$\delta_i^j(h_t) = \begin{cases} 1, \text{ if } h_t \text{ is such that } f_t(i) = j \\ 0, \text{ otherwise} \end{cases}.$$

Then, firm j's expected profits (the density) from worker i at history h_t are given by the following expression:

$$\pi_i^j(h_t) = \mathbf{E}\left[\sum_{s=0}^{T-t} \left(\delta_i^j(h_{t+s}) y_{t+s}^i - w_{t+s,i}^j(h_{t+s})\right) \mid h_t\right].$$

$y_t^i = 1$). For all workers i such that $q_t^i = 1$, $\rho_t = 1$ is an optimal choice.

Let now consider the ex-ante utility $V_2(\pi, \lambda)$ and in general $V_t(q, \pi)$, $t \geq 2$. We now show that $V_t(q, \pi) = \mathbf{E}[U_t(h_t) \mid h_{t-1}]$ solves the following optimization problem:

$$V_t(q,\pi) = \max_{w,w_z,s_z,\rho_z,\pi_z} pq\{\rho_g[u(w_g) + V_{t+1}(\pi_g, 1)] + (1-\rho_g)V_t^m(k - s_g, 1)\}$$
$$+ (1-qp)\{\rho_u[u(w_u) - \rho_g A(\widetilde{w}(q)) + V_{t+1}(\pi_u, q_u)] + (1-\rho_u)V_t^m(k - s_u, q_u)\} \qquad \text{(Problem 1)}$$

subject to

$$\rho_z \in \{0,1\},\ z = u, g \qquad [\rho]$$
$$q_u = \frac{q(1-p)}{q(1-p)+(1-q)} \qquad [q]$$
$$\widetilde{w}(q) = w_g - w_u \qquad [\alpha]$$
$$\rho_u[u(w_u) - \rho_g A(\widetilde{w}(q)) + V_{t+1}(\pi_u, q_u) - V_t^m(k, q_u)] \geq 0 \qquad [u]$$
$$\rho_g[u(w_g) + V_{t+1}(\pi_g, 1) - V_t^m(k, 1)] \geq 0 \qquad [g]$$

and the budget constraint:

$$pq[\rho_g(p - w_g + \pi_g) + (1-\rho_g)(-s_g)]$$
$$+ (1-pq)[\rho_u(pq_u - w_u + \pi_u) + (1-\rho_u)(-s_u)] \geq \pi. \qquad [\pi]$$

Given a solution to this optimization problem, we construct an equilibrium $[\mathcal{W}, \mathcal{F}, \mathbf{U}, \mathbf{w}, \bar{\mathbf{w}}]$.

Consider a history h_t such that $q_t = 1$ for some worker i (good type). Worker i's ex-post utility is $U(h_t) = u(w_g) + V_{t+1}(\pi_g, 1)$ when $\rho_{g,t} = 1$, and thus $f_t(i) = f_{t-1}(i)$. It is $U(h_t) = V_t^m(k - s_g, 1)$ when $\rho_{g,t} = 0$, and thus $f_t(i) \neq f_{t-1}(i)$. Ex-post utilities for workers with $q_t \neq 1$ (unknown type) are defined similarly. The equilibrium payments $w_t(h_t)$ corresponds to w_g (resp. w_u) if in h_t we have $q_t = 1$ (resp. $q_t \neq 1$) regardless of the specific firm entry f_t. And the profits values are the expected profits of firm f_t at node h_t regardless of the details of the other firms' offers.

Clearly, when $\rho_z = 1$, $z = u, g$ the constraints $[u]$ and $[g]$ must be satisfied by the equilibrium value of utility $U(h_t)$, since the worker maximizes at each such node. The market values V_t^m satisfies (6) for the same reasons given in the initial period. Notice, indeed, that when $\rho_z = 1$, firms other than f_t make zero payments at equilibrium.

Notice first that if the worker were never to leave after t (that is, $\rho_{g,s} = \rho_{u,s} = 1$ for all $s \geq t$, then the payments will always be made by only one firm and the equivalence between Problem 1 and the equilibrium follows from standard arguments in the recursive contracts literature.[35]

[35] The problem is then a simple extension of Thomas and Worral (1988). See also Ljungqvist and Sargent (2000).

Let's consider now the case where the worker changes firm at t for the first time, say $\rho_{u,t} = 0$ (the case $\rho_{g,t} = 0$ is similar). Since $\rho_{g,t} = 1$,[36] by the monotonicity of V_t^m in its first argument we have $s_g = 0$. Hence, lifetime utility of type g workers solves the recursive problem by the argument made above for market and period 1 values.

Consider now the problem related to w_u, π_u and s_u, when $\rho_u = 0$. We must show that the total wage $w_t(h_t)$ received in equilibrium by a worker i with history $h_t = h_{t-1} \cup (0, q_u, f_t)$ can include some payment $w_t^j(h_t)$ from firm j even though $f_t(i) \neq j$. We further show that the unidimensional choice of s_u suffices to fully describe such payment.

Notice that firms do not care about the timing of payments. In particular, s_u may correspond to a lump sum payment or to a stream of payments during multiple periods. However, since firms take as given existing offers when making new ones, these new offers must complete optimally the pre-existing payments in equilibrium. This implies that if $V_t^m(k, q_u)$ defines the maximal utility the agent can get from a market offer, then $V_t^m(k - s_u, q_u)$ must be the lifetime utility the worker can get given that firm f_{t-1} pays s_u in expected terms, independently of the form of such payments. In equilibrium s_u must hence be optimally chosen. If the stream of payments were not chosen to solve Problem 1, at the moment when firm f_{t-1} made the offer it would be the possible to offer a better contract to the agent. This alternative contract would typically deliver more insurance to the worker.

Notice that a solution to the recursive problem exists since all objective functions are continuous and wage payments can be bounded below by 0 and above by 1. Profits can be bounded above by T and below by $-T$, so as to have a compact choice set. We now use the optimal policy to construct an equilibrium of the game and to show simultaneously existence and the equivalence result.

The functions $w_t(\cdot)$, $\bar{w}_t(\cdot)$, and $U_t(\cdot)$ can be derived directly from the recursive formulation. By construction, the profit values derived by the policies are nonnegative ex-ante. We now need to specify $\mathcal{W}$, $\mathcal{F}$. The proposed equilibrium starts with all workers equally distributed among M_1 firms. If at some date t for some type z we have $\rho_{z,t} = 0$ then assume that all leaving workers get distributed equally among new firms M_t and so on. This equilibrium assignment can be generated by a $\mathcal{W}$ where all firms in a given period offer exactly the same contract, which specifies zero payments for all nodes emanating from an initial node with $f_t \neq j$, and the appropriate distribution of policies $\mathcal{F}$ so as to have an equal distribution among firms. Finally, we complement the recursive policies by setting payments to zero at all node not reached in equilibrium. It is then easy to see that such offers constitute an

[36] We will see that $\rho_{g,t} = \rho_{u,t} = 0$ is never optimal as long as $k > 0$.

equilibrium.

We now show that this problem actually takes the form of the recursive formulation in Section 3 by showing that through the severance payments firms internalize the wage losses (due to the moving cost k) of the worker in case of a transition between firms. They hence operate as planners, solving a constrained efficient allocation problem.

Severance payments in the two-period case We consider first the case where $T = 2$. The equilibrium contract must maximize the agent's equilibrium utility:

$$\max_{w,w_z,s_z,\rho^z} u(w) + pq\left[\rho_g u(w_g) + (1-\rho_g)\, u\left(p - k + s_g\right)\right] \\ + (1-qp)\left\{\rho_u\left[u(w_u) - \rho_g A\left(\widetilde{w}(q)\right)\right] + (1-\rho_u)\left[u\left(pq_u - k + s_u\right)\right]\right\}$$

where the social concerns element is multiplied by ρ_g since when $\rho_g = 0$ they disappear as the pool inside the same firm is homogeneous. The constraints are

$$\begin{aligned} \rho_z \in \{0,1\}\,,\ z = u, g \qquad & [\rho] \\ q_u = \tfrac{q(1-p)}{q(1-p)+(1-q)} \qquad & [q] \\ \widetilde{w}(q) = w_g - w_u \qquad & [\alpha] \\ \left[u(w_u) - A\left(\widetilde{w}(q)\right) - u\left(pq_u - k\right)\right]\rho_u \geq 0 \qquad & [u] \\ \rho_g\left[u(w_g) - u(p-k)\right] \geq 0 \qquad & [g] \end{aligned}$$

Notice that the participation constraint must be satisfied only when $\rho_z = 1$ and that in this case the right hand side is such that there are no payments in the (of-the equilibrium) case the worker left.

The budget constraint is:

$$\begin{aligned} & pq - w + pq\left[\rho_g\left(p - w_g\right) + (1-\rho_g)\left(-s_g\right)\right] \\ & (1-pq)\left[\rho_u\left(pq_u - w_u\right) + (1-\rho_u)\left(-s_u\right)\right] \\ & \geq \pi \end{aligned}$$

where it is taken into account that if $\rho_g = 0$ the firm must pay a severance payment. Denote by ϕ_π the associated multiplier.

Now assume for example that $\rho_u = 0$. The optimal severance payment s_u solves

$$u'\left(pq_u - k + s_u^*\right) = \phi_\pi$$

The first order condition for w is

$$u'\left(w^*\right) = \phi_\pi.$$

Hence we have that $s_u^* = w^* - (pq_u - k)$. Similarly, when $\rho_g = 0$ we have that $s_g^* = w^* - (p - k)$. Typically, $(p-k) > w^*$ hence s_g^* would be a tax but the limited commitment of the worker will imply that $s_g^* = 0$.

Notice that when $pq < 1 - pq$ for example, the two period problem can be written as follows

$$V_1(\pi, q) = \max_{w_u, w_g, \rho^g, \rho^u} u(w) + pqu(w_g) + (1 - qp)\left[u(w_u) - \rho A\left(\widetilde{w}(q)\right)\right]$$

subject to

$$\begin{aligned} \rho \in \{0,1\} & \qquad [\rho] \\ q_u = \tfrac{q(1-p)}{q(1-p)+(1-q)} & \qquad [q] \\ \widetilde{w}(q) = w_g - w_u & \qquad [\alpha] \\ u(w_u) - \rho A\left(\widetilde{w}(q)\right) \geq u(pq_u - k) & \qquad [u] \\ u(w_g) \geq u(p - k) & \qquad [g] \\ pq\left(p - w_g\right) + (1 - pq)\left(pq_u - w_u\right) - (1 - \rho)pqk \geq \pi & \qquad [\pi] \end{aligned}$$

where $\rho = \rho_u$.

Severance payments in the general case We now consider the general case. Our aim is to show that Problem 1 is equivalent to the following Problem 2:

$$V_t(\pi, q) = \max_{w_z, \pi_z, \rho_z} \left\{ \begin{array}{c} (1 - pq)\left[u(w_u) - \rho_g \rho_u A\left(\widetilde{w}(q)\right) + V_{t+1}(\pi_u, q_u)\right] \\ +pq\left[u(w_g) + V_{t+1}(\pi_g, 1)\right] \end{array} \right\} \qquad \text{(Problem 2)}$$

subject to

$$\begin{aligned} \rho_u, \rho_g \in \{0,1\} & \qquad [\rho'] \\ q_u = \tfrac{q(1-p)}{q(1-p)+(1-q)} & \qquad [q] \\ \widetilde{w}(q) = w_g - w_u & \qquad [\alpha] \\ u(w_u) - \rho_u \rho_g A(\widetilde{w}(q)) + V_{t+1}(\pi_u, q_u) \geq V_t^m(k, q_u) & \qquad [u'] \\ u(w_g) + V_{t+1}(\pi_g, 1) \geq V_t^m(k, 1) & \qquad [g'] \end{aligned}$$

$$pq\left[p - w_g + \pi_g - (1 - \rho_g)\, k\right] + (1 - pq)\left[pq_u - w_u + \pi_u - (1 - \rho_u)\, k\right] \geq \pi \qquad [\pi']$$

where

$$V_{T-t}^m(k, q) = \max_{w, \pi} \left\{ \begin{array}{cc} & u(w) + V_{t+1}(\pi, q) \\ s.t. & pq - w + \pi - k = 0 \end{array} \right\}$$

Proposition 7 *Let $w^*, w_z^*, \pi_z^*, \rho_z^*$ be the solution to Problem 2. Then (i) w_z^*, π_z^* solve Problem 1 whenever $\rho_z^* = 1$ for $z = u, g$. (ii) If $\rho_u^* = 0$ then the solution to Problem 1 is obtained by* *setting* $-s_u = pq_u - w_u^* + \pi_u^* - k$ *and* $(w, w_u, \pi_u) = (w^*, w_u^*, \pi_u^*)$. *(iii) If* $\rho_g^* = 0$ *then* $s_g = 0$ *and* $(w, w_g, \pi_g) = \left(w^*, w_g^*, \pi_g^*\right)$.

Notice that we haven't contemplated the case where both $\rho_u = \rho_g = 0$. This is so because, as we will show below, this situation never arises at equilibrium.

Proof. (i) is straightforward. To show (ii) notice that once $-s_u = pq_u - w_u^* + \pi_u^* - k$ the budget constraint in Problem 1 coincides with that of Problem 2. Moreover, the objective function both for w and w_g and π_g are identical in the two problems, and the remaining constraints coincide as well. As a consequence, each solution for those variables of the first problem must also be a solution of the second problem. It remains to be shown that the so defined s_u is optimal for Problem 1. We are going to show that it satisfies the first order conditions. Notice that the first order conditions for s_u and w in Problem 1 are

$$\frac{\partial}{\partial \pi} V_t^m (k - s_u, q_u) = u'(w).$$

where from its definition $V^m(k - s_u)$ solves

$$V_t^m(k - s_u, q_u) = \max_{w_u, \pi_u} \left\{ \begin{array}{ll} & u(w_u) + V_{t+1}(\pi_u, q_u) \\ s.t. & pq_u - w_u + \pi_u \geq k - s_u \\ & = pq_u - w_u^* + \pi_u^*. \end{array} \right\}$$

Notice that by construction, the problem above coincides with that faces in Problem 2 when we have chosen w_u^*,π_u^*. Hence its solution w_u, π_u must in particular be such that $w_u = w_u^*$ and $\pi_u = \pi_u^*$. By the envelope theorem $\frac{\partial}{\partial \pi} V_t^m (k - s_u, q_u) = u'(w_u)$ and from the first order conditions in Problem 2 we have $u'(w^*) = u'(w_u^*)$ hence we are done. **Q.E.D.**

Given the expression for V_t^m above, it is straightforward to see that the objective functions of Problem 1 and Problem 2 are identical. We now show that the budget constraints are also identical. By construction, this is trivially true for the budget constraint [π] and [π']. First, when $\rho_z = 1$, then [z] and [z'] are identical, for $z \in \{u, g\}$. Second, suppose that $\rho_z = 0$. In Problem 1, the constraint [z] disappears. In Problem 2, the participation constraints [z'] are trivially satisfied. Indeed, for all $z \in \{u, g\}$, we have:

$$u(w_z) + V_{t+1}(\pi_z, q_z) = V_t^m(k - s_z, q_z) \geq V_t^m(k, q_z),$$

where the last inequality derives from monotonicity of V_t^m and the fact that $s_z \geq 0$.

The ρ decision To get to the final formulation of the main text we need the following lemma.

Lemma 8 *If $k > 0$, in the optimal contract, $(\rho_u, \rho_g) \neq (0, 0)$.*

Proof. Notice that from the objective function and the participation constraints when $\rho_z = 0$ then $\rho_{z' \neq z} = 1$ is weakly optimal. But then from the budget constraint setting $\rho_{z'} = 1$ is strictly optimal as long as $k > 0$. **Q.E.D.**

Lemma 9 *If in the optimal contract $\rho_u + \rho_g = 1$, then if $pq < 1 - pq$ then $(\rho_u, \rho_g) = (1, 0)$, if $pq > 1 - pq$ then $(\rho_u, \rho_g) = (0, 1)$.*

Proof. We saw above that when $\rho_u + \rho_g = 1$ the only difference for the optimal choice is made by the budget constraint. Hence the result comes immediately since when $pq < 1 - pq$ the good type are the less numerous. **Q.E.D.**

These results lead to the expression in the main text.

Appendix B: Proofs of the Remaining Propositions

All proofs that follow will be based on the recursive formulation of the problem and the differentiability of the value function. The equivalence between the sequential problem and its recursive form it has been shown above. The properties of the associated value function V_t are formally shown in Appendix C.

Proof of Proposition 2 (no social preferences). *(i)-(ii)* Are straightforward. *(iii)-(iv)* Are the key results in HH, whose proof fully applies here. Here is the formal proof. In the absence of social concerns, the problem can be written as follows:

$$V_t(\pi, q) = \max_{w_u, w_g, \rho, \pi_u, \pi_g} (1 - pq)\left[u(w_u) + V_{t+1}(\pi_u, q_u)\right] + pq\left[u(w_g) + V_{t+1}(\pi_g, 1)\right]$$

subject to

$$\begin{aligned} \rho \in \{0,1\} \qquad & [\rho] \\ q_u = \tfrac{q(1-p)}{q(1-p)+(1-q)} \qquad & [q] \\ u(w_u) + V_{t+1}(\pi_u, q_u) \geq V_t^m(k, q_u) \qquad & [u] \\ u(w_g) + V_{t+1}(\pi_g, 1) \geq V_t^m(k, 1) \qquad & [g] \\ pq(p - w_g + \pi_g) + (1 - pq)\left(pq_u - w_u + \pi_u\right)) - (1 - \rho)\min\{pq, (1 - pq)\}\, k \geq \pi, \qquad & [\pi] \end{aligned}$$

(i) It is clear from constraint [π] that as long as $k > 0$, $\rho = 1$. *(ii)* When $k = 0$, and again from [π], any retention decision $\rho \in \{0,1\}$ is optimal. As a result, the market assignment of workers to firms is indeterminate. *(iii)* Now take the first order conditions and use the envelope condition to get

$$u'(w_z) = -\frac{\partial}{\partial \pi_z} V_{t+1}(\pi_z, q_z) = \frac{-\frac{\partial}{\partial \pi} V_t(\pi, q)}{1 + \phi_z} \quad \text{for} \quad z = u, g, \tag{7}$$

where ϕ_z is the Lagrange multiplier associated to constraint [z]. Since $\phi_z \geq 0$ wages are weakly increasing as stated in the first part of *(iii)*. To see the second part of the statement, notice that the right hand side in [u] decreases with t. Hence $\phi_{u,t} = 0$ for all t and the result follows from the first order conditions. That is, a constant wage (as required by insurance motives) also solves the participation constraint. *(iv)* When in (7) $\phi_g > 0$, we might have an increase in wage. However, once the type is revealed the problem for these workers becomes stationary. **Q.E.D.**

Proof of Proposition 3 (skill segregation). When $k = 0$ the participation constraint [g] is always binding (otherwise the firm could not make zero ex-ante profits) hence setting $\rho_t = 1$ will induce social concerns. Setting $\rho_t = 0$ increases the objective function and relaxes constraint [u]. **Q.E.D.**

Proof of Proposition 4 (wage compression). In appendix C we show that, despite the fact that the function is not always differentiable or concave, we can without loss of generality restrict attention to differentiable points. We can hence apply the (local-differentiable) Kuhn-Tucker theorem to show existence and non-negativity of the multipliers.[37]

We will focus on interior contracts. The necessary conditions for an interior optimum are (recall that $\frac{\partial A'\left(\widetilde{w}(q_t)\right)}{\partial w_{u,t}} = -A'\left(\widetilde{w}(q_t)\right)$) :

$$u'(w_{u,t}) + \rho_t A'\left(\widetilde{w}(q_t)\right) = \frac{\phi_{\pi,t}}{1+\phi_{u,t}}, \text{ and} \tag{8}$$

$$\frac{\phi_{\pi,t}}{1+\phi_{u,t}} = -\frac{\partial}{\partial \pi_{u,t+1}} V_{t+1}(\pi_{u,t+1}, q_{u,t+1}), \tag{9}$$

From the next period envelope condition we also get

$$-\frac{\partial}{\partial \pi_{u,t+1}} V_{t+1}(\pi_{u,t+1}, q_{u,t+1}) = \phi_{\pi,t+1}. \tag{10}$$

Now since $\widetilde{w}(q_{t+1}) > 0$ and $\rho_{t+1} = 1$. Then, (8)-(9) at $t+1$ and (10) imply:

$$\begin{aligned}\left[u'(w_{u,t+1}) + A'\left(\widetilde{w}(q_{t+1})\right)\right](1+\phi_{u,t+1}) &= -\frac{\partial}{\partial \pi_{u,t+1}} V_{t+1}(\pi_{u,t+1}, q_{u,t+1}) \\ &= u'(w_{u,t}) + \rho_t A'\left(\widetilde{w}(q_t)\right).\end{aligned}$$

Comparing the conditions in two successive periods we have

$$\begin{aligned}&\left[u'(w_{u,t+1}) + A'\left(\widetilde{w}(q_{t+1})\right)\right](1+\phi_{u,t+1}) \\ &= -\frac{\partial}{\partial \pi_{u,t+1}} V\left(q_{u,t+1}, \pi_{u,t+1}\right) \\ &= u'(w_{u,t}) + \rho_t A'\left(\widetilde{w}(q_t)\right).\end{aligned}$$

Since $\widetilde{w}(q_t) < \widetilde{w}(q_{u,t+1})$, by convexity $A'\left(\widetilde{w}(q_{t+1})\right) > A'\left(\widetilde{w}(q_t)\right)$, $\phi_{u,t+1} \geq 0$, and $\rho_t \leq 1$ imply that $u'(w_{u,t+1}) < u'(w_{u,t})$. The result hence follows from the concavity of u. **Q.E.D.**

Proof of Proposition 5 (gradual promotions). The first order conditions in each period for w_g and π_g when $\rho_t = 1$ are (recall that $\frac{\partial A'\left(\widetilde{w}(q_t)\right)}{\partial w_{g,t}} = A'\left(\widetilde{w}(q_t)\right)$

$$\begin{cases} u'(w^0_{g,t})\left(1+\phi_{g,t}\right) - A'\left(\widetilde{w}(q_t)\right) = \phi_{\pi,t} \\ -\frac{\partial}{\partial \pi_{g,t+1}} V_{t+1}(\pi_{g,t+1}, 1)\left(1+\phi_{g,t}\right) = \phi_{\pi,t}, \end{cases} \tag{11}$$

[37] For the technical reader, notice that we are assuming that the Kuhn-Tucker constraint qualifications are satisfied. A sufficient condition for the constraint qualifications is Slater condition for the existence of a strict interior feasible contract.

and the envelope condition next period after a good realization:

$$-\frac{\partial}{\partial \pi_{g,t+1}} V_{t+1}(\pi_{g,t+1}, 1) = \phi_{\pi,t+1}. \tag{12}$$

Equation (11) implies that:

$$u'(w^0_{g,t}) - \frac{1}{1+\phi_{g,t}} A'\left(\widetilde{w}(q_t)\right) = -\frac{\partial}{\partial \pi_{g,t+1}} V_{t+1}(\pi_{g,t+1}, 1).$$

Next, notice that when the type is revealed there is no social concerns since there will be no heterogeneity on the workers. So (11) at $t+1$ after a good realization, and (12) imply that:

$$\begin{aligned} u'(w^1_{g,t+1}) &= -\frac{\partial}{\partial \pi_{g,t+1}} V_{t+1}(\pi_{g,t+1}, 1) \\ &= u'(w^0_{g,t}) - \frac{1}{1+\phi_{g,t}} A'\left(\widetilde{w}(q_t)\right), \end{aligned}$$

and the result follows since when $\widetilde{w}(q_t) > 0$ then $A'\left(\widetilde{w}(q_t)\right) > 0$. **Q.E.D.**

Proof of Proposition 6 (wage non-monotonicity). The fact that $w_{u,t-1} < w_{u,t}$ follow from the proposition on wage compression. We show that $w_{u,t+1} < w_{u,t}$. The first order conditions for workers of unknown type (8)-(9) at periods t imply that:

$$u'(w_{u,t}) + A'\left(\widetilde{w}(q_t)\right) = -\frac{\partial}{\partial \pi_{u,t+1}} V_{t+1}(\pi_{u,t+1}, q_{u,t+1}).$$

The next period envelope condition for an interior contract is:

$$u'(w_{u,t+1}) + \rho_{t+1} A'\left(\widetilde{w}(q_{t+1})\right) = -\frac{\partial}{\partial \pi_{u,t+1}} V_{t+1}(\pi_{u,t+1}, q_{u,t+1}).$$

Hence

$$u'(w_{u,t+1}) + \rho_{t+1} A'\left(\widetilde{w}(q_{t+1})\right) = u'(w_{u,t}) + A'\left(\widetilde{w}(q_t)\right),$$

and the result follows again from $\rho_{t+1} \leq 1$, since $\widetilde{w}(q_t) < \widetilde{w}(q_{t+1})$, the convexity of A and the concavity of u. **Q.E.D.**

Appendix C: Properties of V_{T-t}.

The conditional value functions. For any period t, consider a sequence of history-dependent dummy variables $\boldsymbol{\rho}_t = \{\rho_{t+n}(h_{t+n}) \mid h_t\}_{n=0}^{T-t}$, where h_{t+s} is such that $q_{t+s} \neq 1$ for the agent under consideration (notice that when $q = 1$ we can w.l.o.g. assume $\rho_s(h_s) = 1$). Notice that $\boldsymbol{\rho}_t$ is hence a deterministic vector of length $T - t$. Denote by $\boldsymbol{\Upsilon}_t$ the set of all possible $\boldsymbol{\rho}_t$ and by $V_t(\pi, q, \boldsymbol{\rho}_t)$ the solution to the general recursive problem in Section 3 when the workers' decisions correspond to $\boldsymbol{\rho}_t$. It solves

$$V_t(\pi, q, \boldsymbol{\rho}_t)$$
$$= \max_{w_u, w_g, \pi_u, \pi_g} (1 - pq) \left[u(w_u) - \rho_t A(\widetilde{w}(q)) + V_{t+1}(\pi_u, q_u, \boldsymbol{\rho}_{t+1})\right] + pq \left[u(w_g) + V_{t+1}(\pi_g, 1)\right]$$

subject to $[q], [u], [g]$ and $[\pi]$, where $\boldsymbol{\rho}_{t+1} \in \boldsymbol{\Upsilon}_{t+1}$ is the continuation of $\boldsymbol{\rho}_t$.

Lemma 10 *$V_t(\pi, q, \boldsymbol{\rho}_t)$ is concave and differentiable in π for any t, q and $\boldsymbol{\rho}_t$.*

Proof. By backward induction on the Bellman operator defining V_t starting from $V_{T+1}(\pi, q, \boldsymbol{\rho}_T) \equiv 0$ it can be shown that the conditional function V_t is concave and continuously differentiable in π for all $t, q,\ \boldsymbol{\rho}_t \in \boldsymbol{\Upsilon}_t$. With

$$-\frac{\partial}{\partial \pi} V_t(\pi, q; \boldsymbol{\rho}_t) = u'(w_t) - \rho \frac{d}{dw_t} A(\widetilde{w}(q))$$

This is so since both $u(\cdot)$ and $-A(\cdot)$ are concave and differentiable, by the usual Benveniste and Scheinkman (1979) perturbation argument, each conditional value function V_t is differentiable at any interior point (see also Stokey, Lucas and Prescott (1989); Theorem 4.10). **Q.E.D.**

The upper envelope. We now describe the value function V_t as the upper envelope of the conditional functions $V_t(\pi, q; \boldsymbol{\rho}_t)$ we just defined.

Lemma 11 *The upper envelope function*

$$V_t(\pi, q) = \max_{\boldsymbol{\rho}_t \in \boldsymbol{\Upsilon}_t} V_t(\pi, q, \boldsymbol{\rho}_t) \tag{13}$$

always admits both right and left derivatives that satisfy:

$$V_t^+(\pi, q) \geq V_t^-(\pi, q).$$

Moreover, this upper envelope is almost everywhere differentiable and, whenever the derivative exists, we have:

$$\frac{\partial}{\partial \pi} V_t(\pi, q) = \frac{\partial}{\partial \pi} V_t(\pi, q, \boldsymbol{\rho}_t^*(\pi, q))$$

for all $\boldsymbol{\rho}_t^*(\pi, q)$ maximizers of (13).

Proof. Notice that we can apply directly the Daskin (1967) extended envelope theorem to this problem. The key assumptions of the theorem are that (i) $\frac{\partial}{\partial\pi}V_t(\pi, q; \boldsymbol{\rho}_t)$ must be continuous jointly in $(\pi, \boldsymbol{\rho}_t)$ for all q; and that (ii) the set $\boldsymbol{\Upsilon}_t$ is compact for all t. This is indeed the case since for any given $\boldsymbol{\rho}_t$, $\frac{\partial}{\partial\pi}V_t$ is continuous in π, and (since $T < \infty$) the set $\boldsymbol{\Upsilon}_t$ is a finite set for all t. For a restatement of the theorem, and simple proof, see Lemma 11 in Pavoni (2004). **Q.E.D.**

Lemma 12 *In equilibrium, π will always be chosen so that $V_t(\pi, q)$ is differentiable.*

Proof. We know from the previous Lemma that $V_t^+(\pi, q) \geq V_t^-(\pi, q)$. If we can show that $V_t^+(\pi, q) \leq V_t^-(\pi, q)$, then $V_t^+(\pi, q) = V_t^-(\pi, q)$, and the result follows.

For interior contracts we can distinguish two cases. We restrict to $\rho = 1$ (the case $\rho = 0$ follows *mutatis mutandis*). Notice that by monotonicity the budget constraint [π] is always satisfied with equality.

Case 1. No participation constraint is binding

In this case, social concerns are not active, and the firm in period t maximizes

$$pq\left[u\left(w_g\right) + V_{t+1}\left(\pi_g, 1\right)\right] + (1 - pq)\left[u\left(w_u\right) + V_{t+1}\left(\pi_u, q_u\right)\right]$$

such that

$$pq\left(p - w_g + \pi_g\right) + (1 - pq)\left(pq_u - w_u + \pi_u\right) = \pi.$$

When incentive compatibility constraints are not binding, first-order conditions imply that $p - w_g = \frac{\pi_g}{T-t}$ and $w_g = w_u$. With these expressions, and using the budget constraint, we compute π_g as a function of π_u. The problem thus becomes a free maximization program over the unique variable π_u. At the optimum, the right derivative of the objective function is smaller than the left derivative.

Case 2. Only the good-type workers participation constraint is binding.

We solve for w_g as a function of π_g from [g]. Then, using the budget constraint we are left with two choice variables, π_g and π_u. The optimization program boils down to maximizing

$$pqV_t^m(k, 1) + (1 - pq)\left[u\left(f_1\left(\pi_g, \pi_u\right)\right) - \rho A\left(f_2\left(\pi_g, \pi_u\right)\right) + V_{t+1}\left(\pi_u, q_u\right)\right]$$

The gradient must show a similar inequality and we are again done since both f_1 and f_2 are differentiable since they are composite functions of differentiable functions. **Q.E.D.**

References

[1] Acemoglu, D. (1999), "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence," *American Economic Review* 89, 1259-78.

[2] Acemoglu, D., and J.S. Pischke (1999), "The Structure of Wages and Investment in General Training," *Journal of Political Economy*107, 539-72.

[3] Akerlof, G. and R. Kranton (2000), "Economics and Identity," *Quarterly Journal of Economics* 115, 715-753.

[4] Akerlof, G.A. and J.L. Yellen (1990), "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics* 105, 255-283.

[5] Baker, G., M. Gibbs, and B. Hölmström (1994): "The Internal Economics of the Firm: Evidence from Personnel Data," *Quarterly Journal of Economics* 109, 881-919.

[6] Becker, G. (1957): *The Economics of Discrimination*, Chicago: University of Chicago Press.

[7] Bewley, T.F. (1999): *Why Wages don't Fall during a Recession*, Harvard: Harvard University Press.

[8] Bénabou, R. (1993): "Workings of a City: Location, Education, and Production," *Quarterly Journal of Economics* 108, 619-652.

[9] Bolton, G. (1991): "A Comparative Model of Bargaining: Theory and Evidence," *American Economic Review* 81, 1096-1135.

[10] Bolton, G. and A. Ockenfels (2000): "ERC: A Theory of Equity, Reciprocity and Competition, "*American Economic Review* 90, 166-193.

[11] Bramoullé, Y. (2001): "Interdependent Utilities, Preference Indeterminacy, and Social Networks," mimeo, Université de Toulouse.

[12] Brown, C. and J. Medoff (1989): "The Employer Size Wage Effect," *Journal of Political Economy* 97, 1027-59.

[13] Cabrales, A. and A. Calvó-Armengol: "Social Preferences and Skill Segregation," mimeo, Universitat Pompeu Fabra and Universitat Autònoma de Barcelona.

[14] Cannon, S.A, B. Fallick, M. Lettau, and R. Saks (2001), "Has Compensation Become More Flexible?", *Research in Labor Economics* 20, 243-69.

[15] Charness, G. and M. Rabin (2002): "Understanding Social Preferences with some Simple Tests," *Quarterly Journal of Economics* 117, 817-869.

[16] Daskin J.-M. (1967) *The Theory of Max-Min and its Applications to Weapon Allocation Problems*, New-York: Springer-Verlag.

[17] Davis, S. and J. Haltinwanger (1991): "Wage dispersion between and within US manu-

facturing plants, 1963-1986," *Brookings Paper on Economic Activity: Microeconomics* 115-200.

[18] de Bartolomé, Charles (1990): "Equilibrium and Inefficiency in a Community Model with Peer Group Effects," *Journal of Political Economy.*

[19] Fehr, E. and K. Schmidt (2000a): "A Theory of Fairness, Competition and Cooperation," *Quarterly Journal of Economics* 114, 817-868.

[20] Fehr, E. and K. Schmidt (2000b): "Theories of Fairness and Reciprocity: Evidence and Economic Applications," forthcoming in: M. Dewatripont, L.P. Hansen, and S. Turnovski, *Advances in Economic Theory. Eight World Congress of the Econometric Society*, Cambridge: Cambridge University Press.

[21] Fehr, E., Klein, A. and K. Schmidt (2001): "Fairness, Incentives and Contractual Incompleteness," mimeo, University of Zurich and University of Munich.

[22] Fershtman, C., H.K. Hvide and Y. Weiss (2003): "Cultural Diversity, Status Concerns and the Organization of Work," mimeo, Tel-Aviv University and University of Oslo.

[23] Frank, R.H. (1984): "Are Workers Paid their Marginal Product?," *American Economic Review* 74, 549-571.

[24] Frank, R.H. (1985): *Choosing the Right Pond: Human Behavior and the Quest for Status*, Oxford: Oxford University Press.

[25] Gibbs, D.A. and H. Locking (2000), "Wage Dispersion and Productive Efficiency: Evidence for Sweden," *Journal of Labor Economics* 18, 755-82.

[26] Goldin, C. and R.A. Margo (1992), "The Great Compression: The Wage Structure in the United States at Mid-century," *Quarterly Journal of Economics* 107, 1-34.

[27] Güth, W., Schmittberger, R. and B. Schwarze (1982): "An Experimental Analysis of Ultimatum Bargaining," *Journal of Economic Behavior and Organization* 3, 367-388.

[28] Harris, M. and B. Hölmström (1982), "A Theory of Wage Dynamics," *Review of Economic Studies* 49, 315-333.

[29] Hopenhayn, H. and J.P. Nicolini (1997): "Optimal Unemployment Insurance," *Journal of Political Economy* 105, 412-438.

[30] Katz, L. and K.M. Murphy (1992), "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics* 107, 35-78.

[31] Kramarz, F., Lollivier, S. and L.-P. Pelé (1996): "Wage inequalities and firm-specific compensation policies in France," *Annales d'Economie et de Statistiques* 41/42, 369-386.

[32] Kremer, M. and E. Maskin (1996): "Wage Inequality and Segregation by Skill," NBER Working Paper 5718.

[33] Ledyard, J.O. (1995): "Public Goods: A Survey of Experimental Research," in J.H. Kagel and A.E. Roth eds., *Handbook of Experimental Economics,* Princeton NJ: Princeton University Press.

[34] Legros, P., and A. Newman (2002) "Monotone Matching in Perfect and Imperfect Worlds," *Review of Economic Studies* 69: 925-942

[35] Legros, P., and A. Newman (2004) "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferable Utilities," mimeo, University College London and Université Libre de Bruxelles.

[36] Levine, D. (1998): "Modelling Altruism and Spitefulness in Game Experiments," *Review of Economic Dynamics* 1, 593-622.

[37] Ljungqvist, L., and T. J. Sargent (2000), *Recursive Macroeconomic Theory*, MIT Press.

[38] Mas-Colell, A., Whinston, M.D. and J.R. Green (1995): *Microeconomic Theory*, Oxford University Press.

[39] Pavoni, N. (2004): "Optimal Unemployment Insurance with Human Capital Depreciation and Duration Dependence," mimeo, University College London. http://www.ucl.ac.uk/~uctpnpa

[40] Rabin, M. (1993): "Incorporating Fairness into Game Theory," *American Economic Review* 83, 1281-1302.

[41] Rey-Biel, P. (2002): "Inequity Aversion and Team Incentives," mimeo, University College London.

[42] Rotemberg, J. (1994): "Human Relations in the Workplace," *Journal of Political Economy* 102, 684-717.

[43] Roth, A.E.(1995): "Bargaining Experiments," in J.H. Kagel and A.E. Roth eds., *Handbook of Experimental Economics,* Princeton NJ: Princeton University Press.

[44] Sobel, J. (2000): "Social Preferences and Reciprocity," *mimeo.*

[45] Saint-Paul, G. (2001): "On the Distribution of Income and Worker Assignment under Intra-firm Spillovers, with an Application to Ideas and Networks," *Journal of Political Economy* 109, 1-37.

[46] Schelling, T. (1971): "Dynamic Models of Segregation," *Journal of Mathematical Sociology* 1, 143-186.

[47] Thomas, J. P., and T. Worral (1988) "Self-Enforcing Wage Contracts," *Review of Economic Studies* 55(4): 541-554.