Investment Banking Careers

PRELIMINARY AND INCOMPLETE

ULF AXELSON and PHILIP BOND*

October 31, 2008

ABSTRACT

We set up a general equilibrium model of a labor market where moral hazard problems are a key concern. We show that variation in moral hazard across industries explains cross-sectional patterns in contract terms, work patterns over time, and promotion structures. In particular, we explain why very high-profile jobs such as investment banking pay more and give higher utility to the employee than other jobs, even though agents employed do not have any skill advantage. These jobs are also characterized by high firing rates, and inefficiently long hours spent on mundane tasks early on in the career - the "dog years". We also show that agents who are unlucky early on in their careers, either because they do not land a high-profile job or because they lose a high-profile job, suffer a life-long disadvantage in the labor market even when there are no skill differences between workers. Finally, we show why employers may rationally reject more talented job applicants in favor of less talented ones – Smart workers may be "over qualified", or "too hard to manage", because their relatively high outside options make them respond less to firing incentives.

*Ulf Axelson is with the Stockholm School of Economics and SIFR, and Philip Bond is with the Wharton School, University of Pennsylvania. Address correspondence to Ulf Axelson, SIFR, Drottninggatan 89, 113 60 Stockholm, Sweden, or e-mail: ulf.axelson@sifr.org.

Jobs differ widely in the contract terms they offer in terms of salary, job security, work hours, and promotion possibilities. As an example, think about the career choices facing a graduating MBA student. At one extreme, he can aim to become an investment banker, where pay is quite high, especially in case of promotion. On the other hand, job security is very low, and the initial years are characterized by 100 hour work weeks, many of which are spent on tasks that well-trained secretaries probably could do as efficiently (gathering data, and preparing spreadsheets and power-point presentations). At the other extreme, he can try to get a work in the finance department of an industrial company, working fewer hours, and having much greater job security, albeit at a lower salary.

If given offers from both types of firms, most MBAs chose the investment banking job despite the gruelling stress and work hours involved. In fact, casual observation suggests that tough work conditions are strongly correlated with attractiveness of jobs. And for good reason. In an intriguing paper, Paul Oyer (forthcoming) has shown that the MBA student who ends up being lucky enough to get a Wall Street job has an expected lifetime income that is \$1.5 million to \$5 million higher in present value terms than an equally skilled student who does not. He finds it hard to explain this with extra cost of effort, so it appears to be a real utility gain. This suggests that the initial job allocation is more than just "assortative matching" where more talented workers get assigned to more important jobs and get paid their outside opportunity. Instead, there is an element of lottery where the lucky ones get the big prize on Wall Street.

On top of this, it seems like the initial lottery has life-long effects on careers and lifetime earnings. Oyer shows that if a student is unsuccessful in getting hired to a Wall Street job when he graduates because of bad market conditions rather than lack of interest or lack of ability (for example because it happens to be a down year for Wall Street), he is very unlikely to be able to pursue an investment banking career in later years, even if Wall Street happens to be booming. The effect on lifetime income appears to be too big to be explained by switching costs for the worker. Instead, it seems to simply be hard to enter high-profile sectors unless you do it very early in your career.

This is a more general labor-market phenomenon. Older workers who lose their jobs because of economic contraction in their industry face much worse prospects than young workers, even if the young workers have less experience and talent. In a sense, there is a "stigma of failure" in the labor market that seems to result from bad luck rather than bad characteristics or low effort. There is also a related "stigma of success" – a person who has reached a certain position within one industry or company often gets rejected if he applies for a lower-level position in another firm or industry on the grounds that he is "overqualified".

We try to explain all of the features above within a very natural labor market model where differences in moral hazard problems across industries is a key ingredient. We show that industries where moral hazard problems are bigger are more attractive to workers, because they typically have to be paid more than their reservation utility to behave properly. In general equilibrium, this is a relative statement – for higher-than-average moral hazard industries, the incentive constraint rather than the participation constraint binds, while the opposite is true for lower-than-average moral hazard industries.

The extent to which the incentive constraint is binding will explain the contract terms, promotion structures, firing rates, and composition of work force that an industry uses. Although the jobs in the highest moral hazard industries are always (at least weakly) more attractive to workers in utility terms, employers will use every possible tool to reduce worker rents. The higher the rents, the more the tools will be used. This explains why attractive jobs with high salaries tend to have contract terms that –with the exception of the expected salary– are unappealing to workers, including long hours, and high risk of firing.

Also, wage schedules tend to be very steep in the high moral hazard industries. It is better to reward employees late rather than early in their careers, because this increases the incentives to work hard for promotion. Because of the steep wage schedules in high moral hazard industries, the incentive constraint typically stops binding after promotion. In turn, this means that if the firm has different tasks that need to be performed, it is better to assign the "more important", or higher moral hazard, tasks to promoted workers. The first stages of the career in this high-profile industries will be characterized by very long hours on more menial tasks – "the dog years".

Because high moral hazard industries have a relatively bigger advantage of using dynamic incentives such as threat of firing and promotion schemes, young workers with a longer work life ahead of them will be relatively more attractive to hire for these industries. This explains why it is very hard to enter high-profile industries late in a career. Also, it explains why workers who get fired from high-profile jobs may not be able to find work at a lower level in the same industry – they are "overqualified" or have "too much experience" (a euphemism for being too old).

Most of the analysis we perform is without any skill difference across workers. When we introduce skill differences, some surprising results fall out. In particular, assortative matching does not hold - the most skilled workers do not get the best (highest utility) jobs. This is true if variation of skill is not too big, in which case the effect on productivity of hiring the most skilled worker is not too big. However, the threat of firing has a much smaller bite on the more skilled worker. This is because if he is fired, he will be able to get a job in the best industry for fired workers, a first order effect. In a sense, the skilled worker is "overqualified" or "hard to manage" - he does not respond well to incentives because his outside options are too high.

Related literature: <<TO BE WRITTEN>>

The rest of the paper is organized as follows. In Section I, we describe the setup of the model. In Section II, we show the solution to the static contracting problem, and in Section III we show the solution to the dynamic contracting problem. Section IV discusses the labor market equilibrium, and also extends the analyses to the case of differential tasks within firms and the case of differential skill across workers. Section V concludes.

I. Model Setup

To study the labor market phenomena we are interested in, we need two key elements: Workers of different age, and industries that vary in their degree of moral hazard problems. To this end, we assume a supply $\frac{1}{2}\overline{\lambda}$ of young workers enter the labor market each period, work for two periods, and then exit. Thus, the total supply of workers is $\overline{\lambda}$. Except for age, workers are identical. They all have the same skill, are risk neutral, start out penniless, and have limited liability. (We will analyze a setup where skills differ across workers in Section IV.)

There is a continuum of industries indexed by $k \in [0, \overline{k}]$. A worker employed within an industry a certain period works on a project. Projects vary across industries in the amount

k that is at stake. If the project does not succeed, k is lost, while if the project succeeds, there is a gain G. Workers can reach a probability of success p by incurring a non-pecuniary cost of effort $c\gamma(p)$, where $\gamma(0) = 0$, γ is increasing and convex, $\gamma'(0) = 0$ and $\gamma'(1) = \infty$. The moral hazard element comes from the fact that effort is not observable. The principal in a firm can only observe whether the project is a success or not. As is standard in moral hazard models, and as we will show in more detail below, expected compensation to the worker is therefore larger than his cost of effort.

Our interpretation of the money at stake k deserves some further discussion. For a standard firm that engages in risky projects, it simply refers to the cost or size of a project, and $G + k$ can then be interpreted as the present value of future revenues generated by the project if it succeeds. In the context of an investment bank, we can think of k as what is at stake for the client who hires the investment bank to give advice - for example on a takeover. We can also think of it as assets under management, either for proprietary trading or on behalf of clients.

Suppose the expected wages paid to the worker is $E(w)$. The profit of the firm is then

$$\begin{aligned} & pG - E(w) - (1-p)\,k \\ = \; & pg - E(w) - k, \end{aligned}$$

where $g \equiv G + k$ is the *marginal productivity of labor,* that is, the increase in gross revenues from increasing effort p. We will express everything in terms of g rather than G, as this makes for a cleaner exposition. We assume that the gains G, or equivalently g, from success are decreasing in the number of successes within the industry. Each sector is competitive, so there will be entry until there is no way to make positive profits:

ZERO PROFIT CONDITION: *In equilibrium, at the profit maximizing contract which leads to an expected success probability* p^* *at an expected wage cost of* w^*, *the pay off* g *in the success state is such that firms make zero profit:*

$$p^* g - w^* - k = 0.$$

The following Lemma, which shows that in equilibrium expected revenues, success profitabilities (which is equivalent to effort), and wages are all increasing in k, is almost immediate from the revealed preference and the zero profit condition:

LEMMA 1: *In equilibrium, we must have g strictly increasing in k and w^* and p^* weakly increasing in k.*

Proof: In Appendix.

These simple equilibrium relations are useful for understanding why moral hazard problems are "bigger" in industries where more is at stake. When the possible loss k is bigger, it must be that the gains are also bigger in equilibrium, or else firms would never enter the industry. Since both gains and losses are bigger, it is more important to incentivize the worker to work hard so that the success probability increases. As it turns out, this can only be done by paying the agent more in case of success, which means that even if he does not increase his work effort he will get a higher expected pay. Hence, if a worker ends up in a high k industry, he will typically get both higher wages and higher utility.

We could have modelled the magnitude of moral hazard problems within an industry in other ways without changing the general message of the paper. For example, instead of varying the money at stake, we could increase the noise between unobservable effort and observable outcome, or we could increase the cost of effort. The important feature, which we will come to further down, is the relative extent to which the incentive constraint rather than the participation constraint binds across industries.

For use below, we also make the following fairly innocuous assumption on the shape of the effort cost:

ASSUMPTION 1: $p\frac{\gamma'''(p)}{\gamma''(p)} > -1$ and $\lim_{p\to 0} p\frac{\gamma'''(p)}{\gamma''(p)} < \infty$.

A. Contracting environment

There are two classes of contracts: Contracts with old workers, who only work for one period, and contracts with young workers, who might work for two periods with the same employer unless he is fired or leaves voluntarily after the first period. For old workers, a contract is simply a fixed payment w and an extra payment Δ in case of success, where limited liability requires:

LIMITED LIABILITY: $w \geq 0$ and $\Delta \geq -w$.[1].

The maximization problem for the firm is:

$$\max_{w \geq 0, \Delta \geq -w} \pi = p\left(g - \Delta\right) - w - k, \tag{P1}$$

such that

$$p \in \arg\max_{\tilde{p}} \tilde{p}\Delta + w - c\gamma\left(\tilde{p}\right), \tag{IC1}$$

and

$$p\Delta + w - c\gamma\left(p\right) \geq v. \tag{PC1}$$

Here, (IC1) is the incentive compatibility constraint of the worker, ensuring that he indeed finds it optimal to exert effort p, and (PC1) is the participation constraint of the worker that ensures him his reservation utility $v \geq 0$. In equilibrium, the reservation utility v will be set such that the labor market for old workers clear, and the profit g in the success state will be determined by the zero profit condition. For now we keep v and g arbitrary, except that we note that g must be strictly increasing in k. Given v and g, denote the firm profit that solves Program (P1) by $\pi\left(v, g\right)$.

For a firm hiring a young worker, we can describe the contract as a quadruplet $\{w, v_s, v_f, f\}$, where w is a fixed payment in period 1, v_s is the continuation utility promised to the worker in case of success, v_f is the continuation utility promised to the worker in case he fails but is retained, and f is the firing probability in case of failure.[2]

We assume that contracts have to be renegotiation proof, so that they lie on the Pareto frontier in period 2. Given continuation utility v_i, the profit to the firm is given by $\pi\left(v_i, g\right)$ from the solution of Program (P1). We can then express the dynamic maximization problem

[1] This limited liability condition could be relaxed if the old worker had positive wealth to pledge at the beginning of the period. It turns out that he never does in equilibrium.

[2] It turns out never to be optimal to give severance pay, i.e., pay the worker if he is unsuccesful and fired, and never to be optimal to fire the worker if he is succesful, which is why we omit notation for these type of contract features.

for the firm as:

$$\max_{f,v_s\geq \underline{u}(g),v_f\geq \underline{u}(g)} p\left(g+\pi\left(v_s,g\right)\right)+\left(1-p\right)\left(1-f\right)\pi\left(v_f,g\right)-k, \tag{P2}$$

such that:

$$p\in \arg\max_{\tilde{p}} \tilde{p}v_s+\left(1-\tilde{p}\right)\left(\left(1-f\right)v_f+fU\right)-c\gamma\left(p\right), \tag{IC2}$$

and:

$$\max_{\tilde{p}} pv_s+\left(1-p\right)\left(\left(1-f\right)v_f+fU\right)-c\gamma\left(p\right)\geq V. \tag{PC2}$$

Here, (IC2) is the incentive constraint ensuring that the agent works p in period 1, where U is the expected utility for the worker if he gets fired and seeks employment in the industries hiring old workers. The participation constraint (PC2) ensures that worker utility is at least his outside opportunity V, which is solved for in equilibrium. The constraints $v_s, v_f\geq \underline{u}\left(g\right)$ are the renegotiation-proofness conditions, where $\underline{u}\left(g\right)$ denotes the utility the worker gets in the solution to Program (P1) if we set $v=0$ (so that the participation constraint does not bind). Given V, U, and g, denote the firm profits that solves Program (P1) by $\Pi\left(V,U,g\right)$.

B. Equilibrium

Denote the set of old and young workers in industry k by λ_{ko} and λ_{ky}, respectively. It is clear that if the participation constraints in the programs above are ignored, a firm is always at least as well off hiring a young worker as hiring an old worker. This is so since one possibility in the dynamic program is to use a repeated version of the one-period contract that solves the static program. We will see that firms using young workers in fact always do better than this. Hence, young workers are more attractive in the labor market. Since high-k industries turn out to be more attractive to workers, young workers will match with high-k industries. We will guess and verify that in equilibrium, there is a $\widehat{k}\in\left(0,\bar{k}\right)$ such that industries with $k<\widehat{k}$ only hire old workers, whereas industries with $k\geq\widehat{k}$ only hire young workers (some of which may be retained when they are old). Note that this implies that we can have $\lambda_{ko}>0$ for $k\geq\widehat{k}$, but must have $\lambda_{ky}=0$ for $k<\widehat{k}$. Furthermore, all old workers in industries $k<\widehat{k}$ must be workers who got fired from a $k\geq\widehat{k}$ industry after the first period, and no fired worker gets rehired into an industry with $k\geq\widehat{k}$.

Given this conjecture, we define an equilibrium in the labor markets as a set $\{v, V, U, g(k), \lambda_{ko}, \lambda_{ky}\}$ satisfying the following conditions:

- *Firms earn zero profits:* $\Pi(V, U, g(k)) = 0$ for $k \geq \widehat{k}$ and $\pi(v, g(k)) = 0$ for $k < \widehat{k}$.
- *No poaching of workers:* The reservation utility V for young workers is such that $\Pi(V, U, g(k)) \leq 0$ for $k < \widehat{k}$, and the reservation utility v is such that $\pi(v, g(k)) \leq 0$ for $k \geq \widehat{k}$.
- *Labor markets clear:*

$$\int_0^{\overline{k}} \lambda_{ko} dk \leq \frac{1}{2}\overline{\lambda},$$
$$\int_{\widehat{k}}^{\overline{k}} \lambda_{ky} dk \leq \frac{1}{2}\overline{\lambda}.$$

- The outside opportunity U in the second period is given by

$$U = \frac{\int_{k<\widehat{k}} \lambda_{ko} u_k dk}{\int_{k<\widehat{k}} \lambda_{ko} u_k dk},$$

  where u_k is the utility earned by an old worker in industry k.

We prove existence of the equilibrium in Appendix B. In the next two sections, we solve the static and dynamic optimization problems.

II. Solving the Old Worker Contracting Problem

As a benchmark, we first solve for the first-best effort p given the pay-off g. This effort $p_{FB}(g)$ simply sets the marginal benefit of increasing the success probability equal to the marginal cost of effort:

$$g = c\gamma'(p_{FB}(g)).$$

Note that p_{FB} is strictly increasing in g, and hence in k. We will see below that the effort level will typically be lower than this because of the moral hazard problems.

The incentive compatibility condition (IC1) can be written as:

$$\Delta = c\gamma'(p).$$

Using this relation, we can write the contract in terms of p and w, and we write the utility the worker gets from the contract by $u(p,w)$:

$$u(p,w) \equiv c(p\gamma'(p) - \gamma(p)) + w.$$

Note that the utility is strictly increasing in both arguments. In particular, the higher the effort p, the higher the utility of the agent. The maximization problem (P1) can now be rewritten as:

$$\max_{w \geq 0, p} pg - u(p,w) - c\gamma(p),$$

such that:

$$u(p,w) \geq v.$$

The solution to this problem depends on the size of v. First, suppose v is so small that the participation constraint is not binding. Then, it is easy to see that it is optimal to not give the agent any fixed pay w, and to set p from the first order condition such that:

$$g = \frac{\partial u(p,0)}{\partial p} + c\gamma'(p), \tag{1}$$

that is, the firm sets the success probability such that the marginal benefit g equals the marginal cost of effort plus the marginal increase in surplus that is captured by the agent. We call the solution to (1) $p_{SB}(g)$ for the second-best effort level. Note that $p_{SB}(g) \in (0,1)$, since the right-hand side of Equation (1) is strictly increasing in p, is 0 at $p = 0$, and goes to infinity as p goes to 1.[3]

It is easy to see that $p_{SB}(g) < p_{FB}(g)$, the first-best level of effort. This is a standard

[3] To see that the right-hand side increases in p, note that the derivative with respect to p is equal to

$$2\gamma''(p) + p\gamma'''(p).$$

From Assumption 1, this is strictly positive.

result in moral hazard models; Because the worker captures some of the surplus, there is an extra cost to the firm of increasing effort, which reduces the optimal level of effort. More important for our purposes is that $p_{SB}(g)$ increases with the amount at stake k, since g is strictly increasing in k. This means that the utility of the agent also increases with k. This is why we call high k industries "attractive" or "high moral hazard industries", since the surplus given to agents is typically higher.

The firm will set $w = 0$ and $p = p_{SB}(g)$ as long as the participation constraint is not binding, that is, as long as $u(p_{SB}(g), 0) \geq v$. Now suppose $u(p_{SB}(g), 0) < v$, so that the participation constraint is binding. There are two ways of increasing the worker's utility to satisfy the participation constraint: Either increase p, or increase w. Increasing p is better for the firm as long as

$$g > c\gamma'(p),$$

that is, as long as p is below the first-best level. If the promised utility v to the agent is so large that the participation constraint is not satisfied even at the first-best effort level, that is, if $u(p_{FB}(g), 0) < v$, it is better to increase agent utility by a fixed payment w instead of increasing the effort p. We collect the solution to the one-period problem in the following lemma:

LEMMA 2: *The solution* $\{p(v,g), w(v,g)\}$ *to the one-period problem is given by:*

$$\begin{aligned}
p(v,g) = p_{SB}(g), w(v,g) = 0 &\quad \text{if } v \leq u(p_{SB}(g), 0), \\
u(p(v,g), 0) = v, w(v,g) = 0 &\quad \text{if } u(p_{FB}(g), 0) \geq v > u(p_{SB}(g), 0), \\
p(v,g) = p_{FB}, u(p_{FB}, w(v,g)) = v &\quad \text{if } v > u(p_{FB}(g), 0).
\end{aligned}$$

For use below, define the one-period profit $\pi(v,g)$ by

$$\pi(v,g) \equiv p(v,g)\, g - u(p(v,g), w(v,g)) - c\gamma(p(v,g)) - k. \tag{2}$$

It is easy to verify that $p(v,g)$ is weakly increasing in both arguments, and that $\pi(v,g)$ is weakly decreasing in v and strictly increasing in g.

A. Equilibrium for industries hiring old workers

We now solve for the equilibrium for industries $k < \hat{k}$ that only hire old workers.

Note that there is one equilibrium reservation utility v for old workers which is the same for all industries that employ only old workers, while the equilibrium g varies across industries. The zero profit condition sets $\pi(v, g) = 0$.

The equilibrium will be such that the participation constraint binds for lower moral hazard industries but not for higher moral hazard industries. Thus, there is some $k^- < \hat{k}$ such that for $k \in \left[k^-, \hat{k}\right)$, the participation constraint does not bind. In that case, we know from Lemma 2 that $w = 0$ and $p = p_{SB}(g)$. Then, from the definitions of $\pi(v, g), p(v, g)$, and u, the zero profit condition becomes

$$\pi(v, g) = p_{SB}g - cp\gamma'(p_{SB}) - k = 0,$$

that is,

$$g = c\gamma'(p_{SB}) + \frac{k}{p_{SB}}. \tag{3}$$

From the definition of p_{SB}, we also have:

$$g = c\gamma'(p_{SB}) + cp_{SB}\gamma''(p_{SB}). \tag{4}$$

From Expressions (3) and (4), the equilibrium effort $p(k)$ for these industries is defined implicitly by:

$$c\gamma''(p(k))\, p^2(k) = k. \tag{5}$$

Since the right-hand side of (5) is increasing in $p(k)$, we have that $p(k)$ increases in k. Hence, so does $u(p(k), 0)$, the utility given to the agent. This verifies that the participation constraint is non-binding for industries with $k \geq k^-$ where k^- is defined by

$$u\left(p\left(k^-\right), 0\right) = v.$$

For industries with $k < k^-$, the participation constraint binds, so $u(p, w) = v$ for these industries. We first show that in equilibrium, it cannot be the case that $w > 0$ for these industries:

LEMMA 3: *In equilibrium, the fixed payment w must be zero in industries hiring only old workers.*

Proof: In Appendix.

Lemma 3 implies that for industries such that $k < k^{-}$, we must have $u(p, 0) = v$. Hence, the equilibrium effort p is $p(k^{-})$ as defined above, so $p(g, v)$ is constant across industries in this interval. The equilibrium success payoff g is given from the zero profit condition as

$$g = c\gamma'(p) + \frac{1-p}{p}k.$$

We collect the solution to the equilibrium in the following Proposition:

PROPOSITION 1: *The labor market equilibrium for old workers consists of employers with $k < \hat{k}$ for some $\hat{k} < \overline{k}$ and a reservation utility v for old workers that clears the labor market. No employer pays a fixed wage w. There is a $k^{-} < \overline{k}$ such that for employers with $k \in \left[k^{-}, \hat{k}\right)$, the equilibrium effort $p(k)$ is given by*

$$c\gamma''(p(k))\, p^2(k) = k,$$

where k^{-}solves

$$u\left(p\left(k^{-}\right), 0\right) = v.$$

For $k \leq k_1$, we have $p = p(k_1)$. The equilibrium effort and agent utility is increasing in k, and strictly so for $k \in [k_1, k^*]$.

Proposition 1 shows a few of the general properties of contracts that we stress in the paper. The incentive constraint binds for firms with bigger moral hazard problems (money at stake), while the participation constraint binds for firms with lower moral hazard problems. It is better to end up in one of the "high-profile" industries, since they give workers higher utility. On the other hand, you work more in these industries, but this is not enough to outweigh the higher pay. Hence, the labor market is a lottery, with some workers being luckier than others.

As we will see, the old workers in the industries described in Proposition 1 are all worse off than even the unluckiest young workers. We now describe contracts for young workers.

III. Solving the Young Worker Contracting Problem

The zero profit condition is now given by

$$p\left(g+\pi\left(v_s,g\right)\right)+\left(1-p\right)\left(1-f\right)\pi\left(v_f,g\right)-k=0. \tag{ZP}$$

Rewriting the incentive compatibility condition (IC2) as the first order condition of the worker's maximization problem, we have the first period effort p defined implicitly by:

$$v_s-\left(\left(1-f\right)v_f+fU\right)=c\gamma'\left(p\right). \tag{IC3}$$

Dynamic contracts allow the firm to use some extra tools for eliciting effort by the worker. First, the reward (continuation utility) the worker gets after success can be paid out partly by allowing the worker to work on the moral hazard task, which is a gain for the employer - the second period effort comes "for free". Second, he can choose to fire the worker after failure to increase the incentives to work in period 1.

Lemma 4 below shows that it is always optimal to "promote" the worker so that he has more access to the moral hazard task (higher p) after success. Also, if the worker fails in the first period but is retained, he is given less access to the moral hazard task (lower p).

LEMMA 4: *The amount of work $p\left(v_s,g\right)$ after success is strictly bigger than both work p in the first period and work $p\left(v_f,g\right)$ after failure.*

Proof: In Appendix.

Note that Lemma 4 shows that effort goes *up* over the career as the worker gets promoted. We would like to downplay the actual increase in work hours at this stage; rather, we want to stress the fact that promotion leads to more work on the important task, which in turn gives the worker high rents. We will see in Section VI that when there are more tasks to be performed within an organization, the worker will typically work longer hours early on in the career but on more "menial" (lower moral hazard) tasks. As he gets promoted, he works less hours, but all on the important task.

We now go on to show how the solution to the dynamic contracting problem varies across industries with differing k. We have already shown in Lemma 1 that average work and wages increase (weakly) with k. We now show that the firing probability f also increases with k, but in compensation for working harder and having a higher risk of getting fired, workers also get higher utility in higher k industries.

Because utility increases in k, there will be some $k^+ \leq \overline{k}$ such that for $k \geq k^+$, the participation condition does not bind, while for $k \in \left[\hat{k}, k^+\right)$, the participation condition binds. We now characterize the solution over these two intervals in turn.

A. Non-binding participation constraint for young workers: The $k \geq k^+$ case.

When the participation constraint is not binding, we start by showing that the worker is always fired after failure in the first period:

LEMMA 5: *The firing rate after failure must be $f = 1$ if the participation constraint is not binding.*

Proof: Let (v_s, v_f, f) be a contract that maximizes the firm's profits in equilibrium, with g the associated equilibrium profit after success. Consequently, both $\pi(v_s, g)$ and $\pi(v_f, g)$ must be non-positive, since otherwise the firm could make strictly positive profits by using a repeated one-period contract. So the zero profit condition (ZP) implies that

$$p\left(g + \pi\left(v_s, g\right)\right) \geq k > 0.$$

Suppose contrary to the claim that $f < 1$. We derive a contradiction by showing that profits are strictly greater if the firm raises f to 1, while leaving v_s and v_f unchanged. Write p^f and p^1 for the original and new values of the success probability p, and note that $p^f < p^1$

(since $v_f > U$). The change in the firm's profit is

$$\begin{aligned} & gp^1 + p^1\pi(v_s, g) - gp^f - p^f\pi(v_s, g) - \left(1 - p^f\right)(1 - f)\,\pi(v_f, g) \\ = & \left(p^1 - p^f\right)(g + \pi(v_s, g)) - \left(1 - p^f\right)(1 - f)\,\pi(v_f, g)\,. \end{aligned}$$

By the observations above this is strictly positive. ■

Using Lemma 5, the maximization problem when the participation constraint is not binding is given by:

$$\max_{v_s \geq u(p_{SB}(g),0)} p\,(g + \pi(v_s, g))\,,$$

such that p solves:

$$v_s - U = c\gamma'(p)\,. \tag{IC4}$$

Here, (IC4) is the incentive compatibility condition expressed as a first order condition.

The utility of the agent given v_s is:

$$u(v_s) = pc\gamma'(p) + U,$$

and the zero profit condition is given by:

$$p\,(g + \pi(v_s, g)) = k.$$

We now show that over the interval $k \in \left[k^+, \overline{k}\right]$, work in both the first and second period as well as worker utility are strictly increasing.

LEMMA 6: *Suppose the participation constraint in the dynamic problem is non-binding. Then, the first-period effort p, the second-period effort after success $p(v_s, g)$ and the utility of the agent are all strictly increasing in k, and $p^2c\gamma''(p) \geq k$ (the worker works more than in a static equilibrium).*

B. Binding participation constraint for young workers: The $k < k^+$ case.

Over this interval, agent utility is constant at V since the participation constraint binds. This is the lowest utility a young worker gets. We first show that per period, this utility is strictly higher than the highest utility an old fired worker gets.

LEMMA 7: *The per-period utility* $\frac{V}{2}$ *for the unluckiest young worker is strictly higher than the utility for the luckiest fired old worker.*

Proof: In Appendix.

We have already argued that a benefit of being young when seeking employment is that you are more likely to end up in high-profile industries, since it is easier to solve moral hazard problems with young workers. Lemma 7 shows that there is an added benefit: Even within an industry with the same k, a young worker can be given a higher utility than an old worker while keeping the firm at zero profits. This is because the optimal work in the static solution is below the first best, so there is room for Pareto improvements if the moral hazard problem can be overcome.

We now show that it can happen that $f < 1$ for this case; That is, the worker is sometimes retained even after failure.

LEMMA 8: *For* $\hat{k}$ *small enough, we have* $f < 1$ *for the marginal industry hiring young workers.*

Proof: In Appendix.

We collect the most important features of the solution to the dynamic problem in the following proposition:

PROPOSITION 2: *In equilibrium, the expected utility of young workers in industries* $k \in \left[\hat{k}, k^+\right)$ *is* V*, where* $\frac{V}{2}$ *is strictly higher than the utility of the luckiest fired worker. The expected utility of young workers in industries* $\left[k^+, \bar{k}\right]$ *is strictly increasing in* k*. Average work per period and firing probability is increasing in* k*, with* $f = 1$ *for* $k \in \left[k^+, \bar{k}\right]$.

Propositions 1 and 2 together imply that the labor markets for both old and young workers are lotteries, where the lucky workers end up in high moral hazard sectors and earn higher rents. This is despite the fact that work conditions are worse the higher k is: Effort is higher and promotion less certain. Furthermore, old fired workers are excluded from the high-profile labor market, and earn strictly lower rents than even the unluckiest young workers. This is because of the bonding benefits of being young: A young worker can pledge his future moral hazard rents as collateral, thereby reducing moral hazard problems in earlier periods. This makes the young worker more efficient in any industry, but makes him extra attractive to high moral hazard industries.

IV. Labor Market Equilibrium: Features and Extensions

In Appendix B, we prove existence of the equilibrium. Here, we show a number of general equilibrium features, as well as results from straight-forward extensions.

A. Lucky cohorts: Temporary industry shocks have life-long effects

Oyer (forthcoming) shows that temporary shocks to Wall Street that affect the number of workers hired in a year have big and life-long effects on the careers of the MBA students who are on the margin of getting hired to an investment bank. Relative to an MBA student who gets an investment banking job, an otherwise identical student who doesn't because he is unlucky enough to graduate in a year when Wall Street is down has a loss of life-time income of up to 5 million dollars in present value terms. He is also very unlikely to ever be able to go into an investment banking career later in life, even if Wall Street is booming. Oyer finds it hard to explain this with differences in skill or preferences. Instead, there seem to be a large element of randomness in who ends up on Wall Street and who does not. Oyer that the difference in income is not a skill premium but rather a compensating differential for the hours, risk, travel, and other factors that go with working on Wall Street.

Our model provides an explanation both to the wage differential, the importance of initial conditions, and the stickiness of careers documented by Oyer, without appeal to either skill differences, development of specific human capital, or other switching costs. Imagine a temporary shock in the demand function for services in the top moral hazard industry ($k = \bar{k}$) in our model, which leads to one less worker being hired. This worker, who instead ends up in a random industry in $\left[\hat{k}, \bar{k}\right)$, can expect a significantly lower life-time income.

Furthermore, his chance to get into a higher-profile industry is gone – as he gets older, he will either stay in his industry or move to a lower k industry. This is because he gets relatively unattractive to high moral hazard industries as he gets older, because he is harder to incentivize.

Consistent with Oyer's findings, this worker also avoids the long hours and risks associated with the top moral hazard industry (where the firing probability is one in case of failure). However, it is not the case that the high pay is set as a compensating differential for the gruelling work conditions. Instead, the causation goes the other way – the fact that there are so much rents to be made causes employers to create work conditions that partly eat up some of those rents. The job is still attractive, though – not only is the life-time income substantially higher, but the life-time *utility* is as well.

B. Multiple tasks: Dog years and promotion

In Lemma 4 we showed that promotion in the model we have set out leads to more work - p goes up. This is attractive to the worker, as he earns higher rents when he works more. It is important to keep in mind that the extra work is on an *important* task, that is, one where the marginal productivity of labor is very high and the moral hazard rents are correspondingly high.

Now imagine that there is an extra tasks, which we call the *menial* task, which can also be performed in the organization. For example, this could involve gathering data, preparing spreadsheets, copying papers, or fetching burgers for the partners. The menial task is also easily monitored. Suppose that if the worker puts effort m on the menial task within a period, the cost of effort is $c\gamma(m)$ just as before, but m is directly observable. For simplicity, let us assume that m leads to the production of services that can be sold at price μm, where μ is decreasing in the supply of the services and set such that all firms earn zero profit in equilibrium. We also assume that in a period, either the menial or the hard task can be performed, but not both at the same time.

B.1. Contracts with old workers

An old worker can either work at the menial or hard task. We have solved for the hard task contract in Section III above. For the menial task, the problem is easy:

$$\max_{m,w\geq 0} \mu m - w$$

such that:

$$w - c\gamma(m) = v.$$

This is solved by setting agent effort at the first-best level:

$$c\gamma'(m_{FB}) = \mu,$$

and setting the wage such that the participation constraint is satisfied:

$$w = v + c\gamma(m_{FB}).$$

In sectors where only old workers are employed, if there is also an unskilled labor force with lower (zero) reservation value that is capable of performing the menial task, no old worker will do the menial task. Furthermore, in the sectors employing young workers, if the menial task is assigned in the second period effort will be first best.

B.2. Contracts with young workers

Suppose the menial task is performed in the second period. Then, the contracting problem essentially reduces to the static problem. In the first period, the agent must be given the static contract where the reward is the wage specified above. Note that this is strictly worse than utilizing the agent for the hard task both periods and hiring old workers (or unskilled workers) to perform the menial task. Thus, we can restrict attention to contracts where the young worker works on the menial task in the first period. The contracting problem with a young worker now has an assigned amount m of the menial task for the worker to perform in the first period. If he does not perform the task, he gets nothing, whereas otherwise he gets the same type of $\{\Delta, w\}$ contract as before. The contracting problem is then:

$$\max_{w\geq 0, \Delta\geq -w, m} \pi = p(g - \Delta) + \mu m - w - k, \tag{P1b}$$

such that

$$p \in \arg\max_{\tilde{p}} \tilde{p}\Delta + w - c\gamma(\tilde{p}), \tag{IC1b}$$

and

$$p\Delta + w - c\gamma(p) - c\gamma(m) \geq v. \tag{PC1b}$$

Again, we have Δ given by the first order condition $\Delta = c\gamma'(p)$. It is also easy to see that now, the participation constraint is always binding – otherwise, increase m, which increases profits. Using this, we can rewrite the maximization problem as

$$\max_{p,m} pg + \mu m - c\gamma(p) - c\gamma(m)$$

such that the limited liability condition $w > 0$ holds, which can be written as:

$$v \geq pc\gamma'(p) - c\gamma(p) - c\gamma(m). \tag{LL}$$

It is easy to see that if the limited liability condition (LL) is satisfied at the first-best levels of effort p_{FB} and m_{FB}, this must be the solution. This is the case if v is bigger than or equal to agent utility at the first-best effort levels with no payment, given by:

$$u(p_{FB}, m_{FB}, 0) = p_{FB}c\gamma'(p_{FB}) - c\gamma(p_{FB}) - c\gamma(m_{FB}),$$

where p_{FB} and m_{FB} are given by:

$$\begin{aligned} \mu &= c\gamma'(m_{FB}), \\ g &= c\gamma'(p_{FB}). \end{aligned}$$

Next, suppose $v < u(p_{FB}, m_{FB}, 0)$. Then, we set $w = 0$. Suppose we increase m and increase p to keep utility constant:

$$\frac{\partial p}{\partial m} = \frac{\gamma'(m)}{p\gamma''(p)}.$$

The first order condition of the profit function by such a change is given by:

$$g = c\gamma'(p) + pc\gamma''(p) - p\gamma''(p)\frac{\mu}{\gamma'(m)}.$$

This shows that p is larger than p_{SB} from the problem without the menial task. Manipulating, we can rewrite the condition above as:

$$g - c\gamma'(p) = p\gamma''(p)\left(c\gamma'(m) - \mu\right),$$

which shows that m is above the first-best level. We also have that:

$$v = pc\gamma'(p) - c\gamma(p) - c\gamma(m),$$

which shows that m is decreasing in v and p is increasing in v.

The take-away from this is that by forcing the worker to work inefficiently much (above the first-best level) on the menial task, the firm can reduce the rents the worker captures on the important task.

The solution varies with k in an interesting way. Both p and m increase with k. That p increases is natural since g is increasing in k, so that the marginal product of labor on the important task is higher for higher k. But work on the menial task, which does not have a higher marginal product of labor for higher k and is already inefficiently high, increases as well. This is to "eat up" the rents the worker earns on the important task.

PROPOSITION 3: *A worker is never assigned the menial task after success. If a young worker is assigned the menial task, he works more than the first best, while his work on the hard task in period 2 is below the first best. Work on both the menial and the hard task increase in k.*

This is our "dog years" result: In high-profile industries, there is typically very long hours early on in the career – more than what is socially efficient, an on less prestigious tasks. As the worker gets promoted, he is rewarded by switching from the menial to the important task.

C. The "hard-to-manage" effect: Being overqualified

So far, we have assumed that all workers are of the same skill and differ only with respect to their age. Within age cohorts, assignment to industries is then a pure lottery. We now break the lottery by introducing differential skill.

In particular we want to show that the type of model we have set up does not necessarily satisfy "assortative matching", that is, that better workers are assigned to better jobs. To this extent, suppose one worker is slightly more skilled than the rest in that he has a slightly lower effort cost.

In a one period economy, this would lead him to be assigned to the industry with the highest moral hazard problems, so he would end up with the best job and earn the highest rents in the labor market. This no longer holds in the dynamic economy. The skilled worker's outside option upon firing is significantly higher than U, the average utility earned in industries below $\hat{k}$. Instead, he will be assured to find work in the best old-worker industry if he gets fired. In turn, this means that high-profile industries that rely on firing incentives will have to give the worker a significantly higher rent to incentivize him to work as much as his peers. Note that this is a first-order effect even when the skill advantage is small, since breaking the lottery in the market for fired workers has a first-order effect. If the skill advantage is small, the direct effect of lower effort cost is not enough to offset the jump in the outside option for the worker. In a sense, the worker is "hard to manage" or "overqualified", and will be turned down by the highest k industries.

PROPOSITION 4: *Assortative matching does not hold.*

Proof: To be completed.

If skill differences become bigger, assortative matching can be restored. In that case, the direct effect of skill on firm productivity can offset the higher outside option.

V. Conclusion

We have set up a general equilibrium labor market model that we think applies particularly well to workers in jobs where the exact link between effort and output is hard to measure. Although we have cast this within an effort model, we think the principles apply to other types of moral hazard as well, such as stealing. We think these problems are extra relevant for the types of jobs sought by MBA students, such as consultancy, investment banking, or general management. We have explained several features of wages, career paths, and contracts in these types of jobs, and how these features covary with the attractiveness of the job. In particular, jobs characterized by higher moral hazard problems where more value is at stake, such as investment banking, will have longer work hours, steeper career paths, higher risk of firing, but also higher compensation. They are also more attractive because in spite of the gruelling work conditions, they give workers higher utility, even when there are no skill differentials between workers.

We have also shown the value of being young for landing high-profile jobs. Being young makes it possible to use future work as collateral, which makes it easier to incentivize the worker early on. They are therefore especially attractive to high moral hazard industries, and if a worker fails to get a job in such an industry early on he will have a very hard time entering later.

Finally, we have shown that extra skill can sometimes be a detrimental asset when applying for these types of jobs. Extra skill makes the worker more attractive to lower moral hazard industries that do not use firing incentives, which means that his outside option when fired is high. This in turn makes it hard to incentivize the worker in the high moral hazard industries – he is considered overqualified, even though he himself would prefer this type of job.

There are several interesting extensions that we have not managed to perform within this paper. One has to do with the boundaries of the firm; our analysis suggests that combining different tasks within one firm and setting up a hierarchy where workers can be moved between more or less important task can improve efficiency. It would be interesting to endogenize the allocation of tasks across firms more fully. A second interesting extension would be to have richer dynamics. In particular, since we have only two periods, we cannot study concepts such as achieving tenure or becoming a partner, that is, be assured of zero firing probability after promotion. We think our model suggests an economic rationale for tenuring workers or making them partners, however. In our model, it is always optimal to postpone rewards for early success to later periods, and then use the promised utility to relax the incentive condition in later periods. After enough success, these promised rewards will be so high that the agent does not need to be incentivized any longer. At that point, we conjecture that it would be optimal to give the worker tenure, as firing incentives no longer are necessary.

Last, it would be interesting to more fully characterize how differential skills would affect our model. We leave this for future research.

Appendix A: Proofs.

Proof of Lemma 1: First, write $w^* (p^*)$ as the minimum expected wage cost for reaching expected success probability p^* in equilibrium. Clearly, $w^* (p^*)$ must be strictly increasing in p^* for any equilibrium p^*, or else a firm is better of choosing a higher p at the same or lower cost. Second, suppose g is not strictly increasing in k, so that there is a $k' > k$ where $g(k') \leq g(k)$. But then, from the zero profit condition, we must have

$$p^* (k') g(k') - w^*(k') - k' = 0,$$

which implies that

$$p^* (k') g(k) - w^*(k') - k > 0.$$

But this is incompatible with equilibrium since a firm with cost k can then make strictly positive profit. Thus, g^* must be strictly increasing in k. Now suppose $p^* (k)$ is somewhere decreasing in k, that is, there is a k, k' such that $k' > k$ and $p^* (k') < p^* (k)$. For this to be an equilibrium, it should not be profitable for either firm to switch contract, that is:

$$p^* (k) g(k') - w^*(k) \leq p^* (k') g(k') - w^*(k'),$$

and

$$p^* (k') g(k) - w^*(k') \leq p^* (k) g(k) - w^*(k).$$

But these together imply that

$$(p^* (k) - p^* (k')) g(k') \leq w^*(k) - w^*(k') \leq (p^* (k) - p^* (k')) g(k).$$

Since and $g(k') > g(k')$ and $p^* (k) > p^* (k')$, this cannot hold. Hence, $p^* (k)$ must be weakly increasing in k, and w^* is strictly increasing in p^*, it must also be weakly increasing k ■

Proof of Lemma 3: Suppose $w > 0$ in equilibrium for some industry, contrary to the claim in the lemma. From Lemma 2, this means that $v > u(p_{FB}, 0)$, that is:

$$v > c\left(p_{FB}\gamma'\left(p_{FB}\right) - \gamma\left(p_{FB}\right)\right),$$

where p_{FB} is defined by:

$$g = c\gamma'\left(p_{FB}\right).$$

This implies that:

$$v > p_{FB}g - c\gamma\left(p_{FB}\right). \tag{6}$$

But the profit function $\pi\left(v, g\right)$ is given by

$$\pi\left(v, g\right) = p_{FB}g - v - c\gamma\left(p_{FB}\right) - k.$$

From Expression (6), this is negative, which is incompatible with the zero profit condition. Hence, we must have $w = 0$ and $v < u\left(p_{FB}, 0\right)$ for all industries. ■

Proof of Lemma 4: To show that $p\left(v_s, g\right) > p$, note that p is given implicitly by (IC3), while $p\left(v_s, g\right)$ is given implicitly by:

$$v_s = c\left(p\left(v_s, g\right)\gamma'\left(p\left(v_s, g\right)\right) - \gamma\left(p\left(v_s, g\right)\right)\right),$$

if $u\left(p_{FB}\left(g\right), 0\right) > v_s$. Rewriting, we have

$$\frac{v_s + c\gamma\left(p\left(v_s, g\right)\right)}{p\left(v_s, g\right)} = c\gamma'\left(p\left(v_s, g\right)\right).$$

Since $\gamma'\left(p\right)$ is increasing in p and since

$$v_s - \left(\left(1 - f\right)v_f + fU\right) < \frac{v_s + c\gamma\left(p\left(v_s, g\right)\right)}{p\left(v_s, g\right)},$$

we have $p < p\left(v_s, g\right)$. If $u\left(p_{FB}\left(g\right), 0\right) < v_s$, we have $p\left(v_s, g\right) = p_{FB}\left(g\right)$, the first-best level, which must always exceed p. This shows that $p < p\left(v_s, g\right)$. To show that $p\left(v_s, g\right) > p\left(v_f, g\right)$, it is enough to show that $v_s > v_f$. Suppose this were not the case, so that $v_s \leq v_f$. Then, an increase of v_s and simultaneous decrease of v_f that keeps agent utility constant has

$$\frac{\partial v_f}{\partial v_s} = -\frac{p}{(1-f)(1-p)}.$$

Such a perturbation changes the profit function by:

$$\frac{\partial p}{\partial v_s}\left(g + \pi(v_s, g) - (1-f)\pi(v_f, g)\right) + p\left(\frac{\partial \pi(v_s, g)}{\partial v_s} - \frac{\partial \pi(v_f, g)}{\partial v_f}\right).$$

This is strictly positive, since $\frac{\partial p}{\partial v_s} > 0$, since $g + \pi(v_s, g) > 0$ from the zero profit condition, since $\pi(v_f, g) \leq 0$, and since

$$\frac{\partial \pi(v_s, g)}{\partial v_s} - \frac{\partial \pi(v_f, g)}{\partial v_f} \geq 0.$$

This last inequality holds since $\pi(v, g)$ is decreasing and concave in v. ■

Proof of Lemma 6: The incentive compatibility condition (IC4) shows that p increases strictly in v_s. Since agent utility is also strictly increasing in v_s, to show that p and agent utility are strictly increasing in k we only need to show that v_s is strictly increasing in k. Suppose this were not the case, so that for $k > k'$ we have $v_s' \leq v_s$. We know $g' > g$. We also know that $p' \leq p$. From revealed preference, it must be the case that:

$$p'\left(g' + \pi(v_s', g')\right) \geq p\left(g' + \pi(v_s, g')\right).$$

But since $p' \leq p$ and $\pi(v_s, g') \geq \pi(v_s', g')$, this can only hold if $p = p'$ and $v_s = v_s'$, that is, the two industries use equivalent contracts. The first order condition with respect to v_s is:

$$\frac{\partial p}{\partial v_s}\left(g + \pi(v_s, g)\right) + p\frac{\partial \pi(v_s, g)}{\partial v_s} = 0.$$

If both industries use the same contract, $\frac{\partial p}{\partial v_s}$,$\pi(v_s, g)$, and p are the same. But for $v_s \geq u(p_{SB}(g), 0)$, we have that $\frac{\partial \pi(v_s,g)}{\partial v_s}$ is (weakly) increasing in g, and since $g' > g$, the two first-order conditions cannot hold at the same time at the same contract. Hence, p and agent utility increase strictly in k. It is easy to see that $p(v_s, g)$ is strictly increasing in either v_s

or g, and since both of these arguments are strictly increasing in k, we have that $p(v_s, g)$ is strictly increasing in k. To show the last part of the lemma, note that the first order condition with respect to v_s is:

$$\frac{\partial p}{\partial v_s}(g + \pi(v_s, g)) + p\frac{\partial \pi(v_s, g)}{\partial v_s} = 0,$$

where $\frac{\partial p}{\partial v_s}$ is given from (IC4) as:

$$\frac{\partial p}{\partial v_s} = \frac{1}{c\gamma''(p)},$$

and $\frac{\partial \pi(v_s,g)}{\partial v_s}$ is given by

$$\begin{aligned} \frac{\partial \pi(v_s, g)}{\partial v_s} &= -1 \text{ if } v_s \geq u(p_{FB}(g), 0), \\ \frac{\partial \pi(v_s, g)}{\partial v_s} &> -1 \text{ if } v_s < u(p_{FB}(g), 0). \end{aligned}$$

Hence, the first order condition implies that

$$\frac{1}{c\gamma''(p)}(g + \pi(v_s, g)) \leq p,$$

or, using the zero profit condition,

$$k \leq p^2 c\gamma''(p),$$

with equality if $v_s \geq u(p_{FB}(g), 0)$. ■

Proof of Lemma 7: The luckiest fired old worker gets utility $u(p_{SB}(g), 0)$ where the price g is set such that a firm with $k = \hat{k}$ earns zero profits with one-period contracts:

$$\pi(u(p_{SB}(g), 0), g) = 0.$$

We now show that if this firm hires a young worker and gives the young worker an expected utility of $2u(p_{SB}(g), 0)$ (so that the per-period utility is the same as for the luckiest old worker), the firm earns strictly positive profits at price g. Therefore, the firm could promise the young worker utility $V > 2u(p_{SB}(g), 0)$ and still break even. This in turn implies that firms hiring young workers in equilibrium must also promise them strictly more than $2u(p_{SB}(g), 0)$ for the no-poaching condition to be satisfied.

We know that the repeated one period contract gives the agent $2u(p_{SB}(g), 0)$. The repeated one-period contract can be implemented by setting $f = 0$, giving the agent the one period contract after both success and failure in the first period, plus an extra promised payment Δ after a first period success where Δ is given by:

$$\Delta = \frac{u(p_{SB}(g), 0) + c\gamma(p_{SB}(g))}{p_{SB}(g)}.$$

Note that the agent also gets $2u(p_{SB}(g), 0)$ by a contract in which $f = 0$, v_s and v_f are given by

$$\begin{aligned} v_s &= u(p_{SB}(g), 0) + \Delta, \\ v_f &= u(p_{SB}(g), 0), \end{aligned}$$

and where the firm gives the worker the optimal one period contract in period 2 given v_s and v_f. This gives the firm profits of giving the agent $2\underline{v}$ can be done by

$$\begin{aligned} \Pi &= p_{SB}(g)(g + \pi(u(p_{SB}(g), 0) + \Delta, g)) + (1 - p_{SB}(g))\pi(u(p_{SB}(g), 0), g) - k \\ &= p_{SB}(g)(g + \pi(u(p_{SB}(g), 0) + \Delta, g)) - k. \end{aligned}$$

The last inequality follows from the zero profit condition. We show that this is positive. We know that

$$p_{SB}(g)(g - \Delta) - k = 0.$$

Thus, we need to show that

$$\pi(u(p_{SB}(g), 0) + \Delta, g) > -\Delta.$$

But we know that

$$\partial \frac{\pi(v,g)}{\partial v} \geq -1$$

for all $v \geq u(p_{SB}(g), 0)$, and

$$\left. \frac{\partial \pi(v,g)}{\partial v} \right|_{v=u(p_{SB}(g),0)} = 0.$$

Hence,

$$\pi(u(p_{SB}(g), 0) + \Delta, g) > \pi(u(p_{SB}(g), 0), g) - \Delta = -\Delta.$$

Thus, the firm earns strictly positive profits while still giving the agent $2u(p_{SB}(g), 0)$. ■

Proof of Lemma 8: Suppose $f = 1$ for the marginal case. The problem of the firm is, taking the price $\underline{g}$ as given from the old worker marginal sector:

$$\underline{g} = c\gamma'(\underline{p}) + \underline{p}c\gamma''(\underline{p})$$

where $\underline{p}$ is defined by

$$k = c\underline{p}^2\gamma''(\underline{p}).$$

The firm's problem is:

$$\max_{v_s} p(v_s)\left(\underline{g} + \pi\left(v_s, \underline{g}\right)\right) - k$$

such that agent participation condition holds condition holds:

$$p(v_s)v_s + (1 - p(v_s))U - c\gamma(p(v_s)) \geq 2\underline{v},$$

where

$$\underline{v} = c\left(\underline{p}\gamma'(\underline{p}) - \gamma(\underline{p})\right),$$

and the incentive condition holds:

$$v_s - U = c\gamma'(p(v_s))$$

The zero profit condition is

$$p(v_s)\left(\underline{g}+\pi\left(v_s,\underline{g}\right)\right)=k.$$

Now, consider a local change in f, with v_s also changed to hold worker utility constant. Worker utility given by:

$$pv_s+(1-p)(1-f)v_f+(1-p)fU-c\gamma(p),$$

where:

$$\frac{v_s-(1-f)v_f-fU}{c}=\gamma'(p),$$

so we have:

$$\frac{dv_s}{df}=\frac{1-p}{p}(v_f-U).$$

The derivative of firm profit with respect to f (moving v_s to hold worker utility constant) is hence:

$$\frac{dp}{df}\left(g+\pi(v_s,g)-(1-f)\pi(v_f,g)\right)+p\pi_v(v_s,g)\frac{1-p}{p}(v_f-U)-(1-p)\pi(v_f,g).$$

At $f=1$, substituting in the zero-profit condition, together with $\pi_v(v_s,g)=-1$ (we show that this holds below) gives

$$\frac{dp}{df}\frac{k}{p}-(1-p)(v_f-U)-(1-p)\pi(v_f,g).$$

Let k_0 denote the marginal case. For this case, there is a v_f such that $\pi(v_f,g)=0$. Note that $\frac{dp}{df}$ is given by:

$$\frac{dp}{df}=\frac{\frac{dv_s}{df}+v_f-U}{c\gamma''(p)}=\frac{v_f-U}{c\gamma''(p)\,p},$$

and since $\frac{dp}{df}>0$, we would like to show:

$$k_0<\gamma''(p)\,cp^2(1-p).$$

We know that:

$$k_0 = c\underline{p}^2\gamma''\left(\underline{p}\right),$$

so we want to show:

$$\underline{p}^2\gamma''\left(\underline{p}\right) < p^2\gamma''(p)(1-p),$$

or:

$$\frac{\underline{p}^2\gamma''\left(\underline{p}\right)}{p^2\gamma''(p)} < 1-p.$$

Suppose $p \to 0$, which is the case when $k \to 0$. .Then, it is enough to show that $\underline{p}/p \leq \frac{1}{1+\lambda}$ for some $\lambda > 0$. At $f = 1$, the worker's utility is:

$$\begin{aligned} & pv_s + (1-p)U - c\gamma(p) \\ = \;& p(v_s - U) - c\gamma(p) + U \\ = \;& p\gamma'(p)c - c\gamma(p) + U. \end{aligned}$$

This has to exceed:

$$2\underline{v} = 2c\left(\underline{p}\gamma'\left(\underline{p}\right) - \gamma\left(\underline{p}\right)\right),$$

so:

$$c(p\gamma'(p) - \gamma(p)) + U \geq 2c\left(\underline{p}\gamma'\left(\underline{p}\right) - \gamma\left(\underline{p}\right)\right).$$

Suppose $U \leq (1-\lambda)\underline{v}$. Then, we have to have

$$(p\gamma'(p) - \gamma(p)) \geq (1+\lambda)\left(\underline{p}\gamma'\left(\underline{p}\right) - \gamma\left(\underline{p}\right)\right),$$

i.e.,

$$\frac{\underline{p}\gamma'\left(\underline{p}\right) - \gamma\left(\underline{p}\right)}{p\gamma'(p) - \gamma(p)} \leq \frac{1}{1+\lambda}.$$

We want to show that as $k \to 0$,

$$\frac{\underline{p}^2\gamma''\left(\underline{p}\right)}{p^2\gamma''(p)} < 1-p.$$

It is sufficient to show that

$$\frac{\underline{p}^2\gamma''\left(\underline{p}\right)}{p^2\gamma''\left(p\right)} \leq \left(1+\frac{\lambda}{2}\right)\frac{\underline{p}\gamma'\left(\underline{p}\right)-\gamma\left(\underline{p}\right)}{p\gamma'\left(p\right)-\gamma\left(p\right)},$$

i.e., that

$$\frac{\underline{p}^2\gamma''\left(\underline{p}\right)}{\underline{p}\gamma'\left(\underline{p}\right)-\gamma\left(\underline{p}\right)} \leq \left(1+\frac{\lambda}{2}\right)\frac{p^2\gamma''\left(p\right)}{p\gamma'\left(p\right)-\gamma\left(p\right)}.$$

The limit of both $\frac{p^2\gamma''(p)}{p\gamma'(p)-\gamma(p)}$ and $\frac{\underline{p}^2\gamma''\left(\underline{p}\right)}{\underline{p}\gamma'\left(\underline{p}\right)-\gamma\left(\underline{p}\right)}$ is

$$2+\lim_{p\to 0}\frac{p\gamma'''}{\gamma''}.$$

So the result follows provided that $\lim_{p\to 0}\frac{p\gamma'''}{\gamma''}$ is finite, which we assume.

We also have to show that as k goes to zero, it is indeed true that $\pi_v\left(v_s,g\right)=-1$ if $f=1$. This amounts to showing that p_g as defined by:

$$c\gamma'\left(\underline{p}\right)+\underline{p}c\gamma''\left(\underline{p}\right)=c\gamma'\left(p_g\right),$$

is smaller than p_v as defined by:

$$c\left(p_v\gamma'\left(p_v\right)-\gamma\left(p_v\right)\right)=v_s,$$

where:

$$v_s-U=c\gamma'\left(p\left(v_s\right)\right).$$

So, in other words, p_v is defined by:

$$c\left(p_v\gamma'\left(p_v\right)-\gamma\left(p_v\right)\right)=c\gamma'\left(p\right)+U.$$

We also know that $p>\underline{p}$. So we have:

$$
\begin{aligned}
c\gamma'(p_g) &= c\gamma'(\underline{p}) + \underline{p}c\gamma''(\underline{p}), \\
c\gamma'(p_v) &= \frac{c\gamma'(p) + c\gamma(p_v) + U}{p_v}.
\end{aligned}
$$

Close to 0, we assume that $\gamma''(\underline{p})$ is bounded away from zero and bounded, and we assume that γ' goes to zero. Dividing one with the other we have

$$
\frac{\gamma'(p_g)}{\gamma'(p_v)} = p_v \frac{\gamma'(\underline{p}) + \underline{p}\gamma''(\underline{p})}{\gamma'(p) + \gamma(p_v) + \frac{U}{c}}.
$$

Suppose contrary to the claim that $p_g > p_v$ so that $\frac{\gamma'(p_g)}{\gamma'(p_v)} > 1$. We then must have:

$$
\frac{\gamma'(\underline{p}) + \underline{p}\gamma''(\underline{p})}{\gamma'(p) + \gamma(p_v) + \frac{U}{c}} \to \infty,
$$

since $p_v \to 0$ as $k \to 0$. But we know $\underline{p} < p < p_v$, so we have:

$$
\frac{\gamma'(\underline{p}) + \underline{p}\gamma''(\underline{p})}{\gamma'(p) + \gamma(p_v) + \frac{U}{c}} < \frac{\gamma'(p) + p\gamma''(p)}{\gamma'(p) + \gamma(p_v) + \frac{U}{c}} < 1 + p\frac{\gamma''(p)}{\gamma'(p)}.
$$

Since we have assumed that $p\frac{\gamma''(p)}{\gamma'(p)}$ is bounded, the result follows. ■

The following lemma shows that there is no fixed payment in period 1 in the dynamic problem:

LEMMA A1: *In equilibrium, the fixed payment w in the first period must be zero in industries hiring only young workers.*

Proof: First, if the participation constraint is not binding, it is obvious that $w = 0$. If the participation constraint is binding, for $w > 0$ to be optimal, we need that a small increase in w and corresponding decrease in v_s that keeps the agent utility constant has no effect on the profit. Agent utility is given by:

$$
pv_s + (1-p)(1-f)v_f + w - c\gamma(p) = V,
$$

so such a permutation that keeps utility constant has:

$$\frac{\partial v_s}{\partial w} = -\frac{1}{p}.$$

The profit function is given by:

$$p\left(g + \pi\left(v_s, g\right)\right) + \left(1 - p\right)\left(1 - f\right)\pi\left(v_f, g\right) - k - w.$$

Setting the derivative of the profit function with respect to w equal to zero gives:

$$\frac{\partial p\left(v_s\right)}{\partial v_s}\left(g + \pi\left(v_s, g\right) - \left(1 - f\right)\pi\left(v_f, g\right)\right) = -p\left(v_s\right).$$

We know that $\pi\left(v_f, g\right) < 0$, so since $\frac{\partial p(v_s)}{\partial v_s} \geq 0$, this implies that $g + \pi\left(v_s, g\right) < 0$. But that is incompatible with the zero profit condition. Hence, $w > 0$ cannot be optimal. ■

Appendix B: Proof of Equilibrium Existence.

<<INCOMPLETE>>

Proof: First, define the total demand for workers in sector k by $\lambda_k = \lambda_{ko} + \lambda_{ky}$. Write $\Lambda_o\left(k^-, \hat{k}\right)$ for surplus demand of old workers:

$$\begin{aligned}
\Lambda_o\left(k^-, \hat{k}\right) &= \int_0^{\bar{k}} \lambda_{ko}\left(k^-, \hat{k}\right) dk - \frac{\lambda}{2} \\
&= \int_0^{\min\left(\hat{k}, k^-\right)} \lambda_{ko} dk + \int_{\min\left(\hat{k}, k^-\right)}^{\hat{k}} \lambda_{ko} dk \\
&\quad + \int_{\hat{k}}^{\max\left(\hat{k}, k^-\right)} \lambda_{ky} dk + \int_{\max\left(\hat{k}, k^-\right)}^{\bar{k}} \lambda_{ky}\left(p\left(k\right) + \left(1 - p\left(k\right)\right)\left(1 - f\left(k\right)\right)\right) dk \\
&\quad - \frac{\lambda}{2}
\end{aligned}$$

Write $\Lambda_y\left(k^-, \hat{k}\right)$ for surplus demand of young workers:

$$\Lambda_y\left(k^-, \hat{k}\right) = \int_{\hat{k}}^{\bar{k}} \lambda_{ky} dk - \frac{\lambda}{2}$$

We make the following assumption:

$$\int_0^{\bar{k}} \lambda_k \left(\phi_{\bar{k}}\right) dk < \frac{\lambda}{2}. \tag{0.1}$$

In words, if all sectors pay $\phi_{\bar{k}}$ (the highest one-period surplus across sectors), demand for workers is insufficient to clear even half the market.

We also assume two INADA conditions: $\lambda_k \left(\phi_k\right) \to \infty$ as $k \to 0$, and $\lambda_k \left(\theta\right) > 0$ for any $\theta < \infty$.

LEMMA B1: *Under Assumption 0.1, there exist k_1^-, k_2^-, $\hat{k}_1$ and $\hat{k}_2$ such that*

$$\begin{aligned} \Lambda_o &> 0 \text{ if } k^- < k_1^- & (0.2) \\ \Lambda_o &< 0 \text{ if } k^- > k_2^- & (0.3) \\ \Lambda_y &> \Lambda_o \text{ if } \hat{k} < \hat{k}_1 & (0.4) \\ \Lambda_y &< \Lambda_o \text{ if } \hat{k} > \hat{k}_2. & (0.5) \end{aligned}$$

Proof: Condition (0.2): From the INADA condition, there exists a $k_1^- > 0$ such that $\Lambda_o \left(k^-, \hat{k}\right) > 0$ for all $\left(k^-, \hat{k}\right) \in \left[0, k_1^-\right] \times \left[0, \bar{k}\right]$. Note that this holds also for $\hat{k} = 0$, since then we have

$$\Lambda_o \left(k^-, 0\right) = \int_0^{k^-} \lambda_{ky} dk + \int_{k^-}^{\bar{k}} \lambda_{ky} \left(p\left(k\right) + \left(1 - p\left(k\right)\right)\left(1 - f\left(k\right)\right)\right) dk - \frac{\lambda}{2},$$

where λ_{ky} blows up as $k \to 0$ and $k^- \to 0$.

Condition (0.3): Follows directly from (0.1).

Condition (0.4): Young workers between sectors $\hat{k}$ and k^- are never fired, and so $\lambda_{ky} = \lambda_{ko}$ in these sectors. In sectors $k > k^-$ a fraction $p\left(k\right) + \left(1 - p\left(k\right)\right)\left(1 - f\left(k\right)\right)$ of workers are retained. So

$$\Lambda_y \left(k^-, \hat{k} = 0\right) - \Lambda_o \left(k^-, \hat{k} = 0\right) = \int_{k^-}^{\bar{k}} \lambda_{ky} \left(1 - p\left(k\right) - \left(1 - p\left(k\right)\right)\left(1 - f\left(k\right)\right)\right) dk.$$

Since $f(k) > 0$ when $k > \max\left\{\hat{k}, k^-\right\}$, it follows that $\Lambda_y\left(k^-, \hat{k} = 0\right) - \Lambda_o\left(k^-, \hat{k} = 0\right) > 0$ for all $k^- < \bar{k}$. The condition follows.

Condition (0.5): At $\hat{k} = \bar{k}$, $\Lambda_y\left(k^-, \hat{k}\right) = -\frac{\lambda}{2} < \Lambda_0\left(k^-, \hat{k}\right)$ for any $k^- \in \left[0, \bar{k}\right]$. Hence there exists $\hat{k}_2 < 1$ such that $\Lambda_y - \Lambda_o < 0$ for all $\hat{k} > \hat{k}_2$. ■

Define

$$\begin{aligned} \overline{\Lambda}_o &\equiv \max_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} \Lambda_o\left(k^-, \hat{k}\right) \\ \overline{\Lambda}_y &\equiv \max_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} \Lambda_y\left(k^-, \hat{k}\right) \end{aligned}$$

We assume $\overline{\Lambda}_o < \infty$ and $\overline{\Lambda}_y < \infty$. (This is a mild assumption about concavity of the profit function, since we have ruled out $k^- = 0$.).

Also, note that $\Lambda_o\left(k^-, \hat{k}\right), \Lambda_y\left(k^-, \hat{k}\right) \geq -\frac{\lambda}{2}$.

Define $g\left(k^-, \hat{k}\right)$ on $\left[\frac{k_1^-}{2}, k_2^- + \frac{\bar{k}-k_2^-}{2}\right] \times \left[\frac{\hat{k}_1}{2}, \bar{k}\right]$ by

$$g\begin{pmatrix} k^- \\ \hat{k} \end{pmatrix} \equiv \begin{pmatrix} k^- \exp\left(\alpha_1 \Lambda_o\left(k^-, \hat{k}\right)\right) \\ \hat{k} \exp\left(\alpha_2\left(\Lambda_y\left(k^-, \hat{k}\right) - \Lambda_o\left(k^-, \hat{k}\right)\right)\right) \end{pmatrix}$$

for some $\alpha_1, \alpha_2 > 0$ to be specified below.

Economically, observe that g is defined so that if there is surplus (insufficient) demand for old workers k^- is raised (lowered); and if demand for old workers exceeds (is less than) demand for young workers, $\hat{k}$ is raised (lowered).

Note that $\Lambda_y\left(k^-, \hat{k}\right) = \Lambda_o\left(k^-, \hat{k}\right) = 0$ at any fixed point of g, since we have defined its domain to exclude the axes $k^- = 0$ and $\hat{k} = 0$. The function g is clearly continuous, so provided we can find $\alpha_1, \alpha_2 > 0$ such that g maps into $\left[\frac{k_1^-}{2}, k_2^- + \frac{\bar{k}-k_2^-}{2}\right] \times \left[\frac{\hat{k}_1}{2}, \bar{k}\right]$, Brouwer's theorem implies the existence of a fixed point, and hence of equilibrium existence.

For the first part of the mapping, we need

$$\max_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) \leq k_2^-+\frac{\left(\bar{k}-k_2^-\right)}{2}$$
$$\min_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) \geq \frac{k_1^-}{2}.$$

Note that for $k^- > k_2^-$, we have from above that $\Lambda_o\left(k^-,\hat{k}\right) < 0$ for any $\left(k^-,\hat{k}\right) \in \left[k_2^-,k_2^-+\frac{1\bar{k}-k_2^-}{2}\right]\times\left[0,\bar{k}\right]$. Thus, $k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) < k^-$ over this interval. Therefore, we only need to check that

$$\max_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) \leq k_2^-+\frac{\left(\bar{k}-k_2^-\right)}{2}.$$

Note that

$$\max_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) \leq k_2^-\exp\left(\alpha_1\overline{\Lambda}_o\right)$$

so it is enough that

$$k_2^-\exp\left(\alpha_1\overline{\Lambda}_o\right) \leq k_2^-+\frac{\left(\bar{k}-k_2^-\right)}{2}$$

for the first condition to hold. For the second condition, note that for $k^- < k_1^-$, $\Lambda_o\left(k^-,\hat{k}\right) > 0$ for any $\left(k^-,\hat{k}\right) \in \left[0,k_1^-\right]\times\left[0,\bar{k}\right]$. Thus, $k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) \geq k^-$ for $k^- \leq k_1^-$. Therefore, we only need to check that

$$\min_{\left(k^-,\hat{k}\right)\in\left[k_1^-,k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) \geq \frac{k_1^-}{2}.$$

Note that

$$\min_{\left(k^-,\hat{k}\right)\in\left[k_1^-,k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\frac{\hat{k}_1}{2},\bar{k}\right]} k^-\exp\left(\alpha_1\Lambda_o\left(k^-,\hat{k}\right)\right) \geq k_1^-\exp\left(-\alpha_1\frac{\lambda}{2}\right).$$

so it is enough that

$$k_1^- \exp\left(-\alpha_1 \frac{\lambda}{2}\right) \geq \frac{k_1^-}{2}$$

for the second condition to hold. Combining, we need to pick α_1 small enough such that

$$\begin{aligned} k_2^- \exp\left(\alpha_1 \overline{\Lambda}_o\right) &\leq k_2^- + \frac{\left(\bar{k} - k_2^-\right)}{2} \\ k_1^- \exp\left(-\alpha_1 \frac{\lambda}{2}\right) &\geq \frac{k_1^-}{2}, \end{aligned}$$

i.e.,

$$\begin{aligned} \alpha_1 \overline{\Lambda}_o &\leq \ln\left(1 + \frac{\left(\bar{k} - k_2^-\right)}{k_2^- 2}\right) \\ -\alpha_1 \frac{\lambda}{2} &\geq \ln \frac{1}{2}. \end{aligned}$$

This is clearly feasible for small enough α_1.

For the second part of the mapping, we need

$$\begin{aligned} \max_{\left(k^-, \hat{k}\right) \in \left[\frac{k_1^-}{2}, k_2^- + \frac{\bar{k} - k_2^-}{2}\right] \times \left[\frac{\hat{k}_1}{2}, \bar{k}\right]} \hat{k} \exp\left(\alpha_2 \left(\Lambda_y\left(k^-, \hat{k}\right) - \Lambda_o\left(k^-, \hat{k}\right)\right)\right) &\leq \bar{k}. \\ \min_{\left(k^-, \hat{k}\right) \in \left[\frac{k_1^-}{2}, k_2^- + \frac{\bar{k} - k_2^-}{2}\right] \times \left[\frac{\hat{k}_1}{2}, \bar{k}\right]} \hat{k} \exp\left(\alpha_2 \left(\Lambda_y\left(k^-, \hat{k}\right) - \Lambda_o\left(k^-, \hat{k}\right)\right)\right) &\geq \frac{\hat{k}_1}{2}. \end{aligned}$$

For $\hat{k} \geq \hat{k}_2$, $\Lambda_y\left(k^-, \hat{k}\right) - \Lambda_o\left(k^-, \hat{k}\right) < 0$. Thus, $\hat{k} \exp\left(\alpha_2\left(\Lambda_y\left(k^-, \hat{k}\right) - \Lambda_o\left(k^-, \hat{k}\right)\right)\right) < \hat{k}$ over this interval. Thus, we only need to check that

$$\max_{\left(k^-, \hat{k}\right) \in \left[\frac{k_1^-}{2}, k_2^- + \frac{\left(\bar{k} - k_2^-\right)}{2}\right] \times \left[\frac{\hat{k}_1}{2}, \hat{k}_2\right]} \hat{k} \exp\left(\alpha_2 \left(\Lambda_y\left(k^-, \hat{k}\right) - \Lambda_o\left(k^-, \hat{k}\right)\right)\right) \leq \bar{k}.$$

Note that

$$\max_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-+\frac{\left(\bar{k}-k_2^-\right)}{2}\right]\times\left[\frac{\hat{k}_1}{2},\hat{k}_2\right]} \hat{k}\exp\left(\alpha_2\left(\Lambda_y\left(k^-,\hat{k}\right)-\Lambda_o\left(k^-,\hat{k}\right)\right)\right)\leq \hat{k}_2\exp\left(\alpha_2\left(\overline{\Lambda}_y+\frac{\lambda}{2}\right)\right)$$

so for the first condition to hold it is enough that

$$\hat{k}_2\exp\left(\alpha_2\left(\overline{\Lambda}_y+\frac{\lambda}{2}\right)\right)\leq \bar{k},$$

i.e.,

$$\alpha_2\left(\overline{\Lambda}_y+\frac{\lambda}{2}\right)\leq \ln\frac{\bar{k}}{\hat{k}_2}.$$

For the second condition, note that for $\hat{k}\leq\hat{k}_1$ and $k^-\leq k_2^-+\frac{\bar{k}-k_2^-}{2}$, we have from above that $\Lambda_y\left(k^-,\hat{k}\right)-\Lambda_o\left(k^-,\hat{k}\right)>0$ over this interval. Hence, $\hat{k}\exp\left(\alpha_2\left(\Lambda_y\left(k^-,\hat{k}\right)-\Lambda_o\left(k^-,\hat{k}\right)\right)\right)>\hat{k}$ for $\hat{k}\leq\hat{k}_1$. Thus, we only need to check that

$$\min_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\hat{k}_1,1\bar{k}\right]} \hat{k}\exp\left(\alpha_2\left(\Lambda_y\left(k^-,\hat{k}\right)-\Lambda_o\left(k^-,\hat{k}\right)\right)\right)\geq\frac{\hat{k}_1}{2}.$$

Note that

$$\min_{\left(k^-,\hat{k}\right)\in\left[\frac{k_1^-}{2},k_2^-+\frac{\bar{k}-k_2^-}{2}\right]\times\left[\hat{k}_1,\bar{k}\right]} \hat{k}\exp\left(\alpha_2\left(\Lambda_y\left(k^-,\hat{k}\right)-\Lambda_o\left(k^-,\hat{k}\right)\right)\right)\geq \hat{k}_1\exp\left(-\alpha_2\left(\frac{\lambda}{2}+\overline{\Lambda}_o\right)\right).$$

Thus, for the second condition to hold, it is enough that

$$\hat{k}_1\exp\left(-\alpha_2\left(\frac{\lambda}{2}+\overline{\Lambda}_o\right)\right)\geq\frac{\hat{k}_1}{2},$$

i.e.,

$$-\alpha_2\left(\frac{\lambda}{2}+\overline{\Lambda}_o\right)\geq\ln\frac{1}{2}.$$

Combining, we need to pick $\alpha_2 > 0$ such that

$$\begin{aligned} \alpha_2 \left(\overline{\Lambda}_y + \frac{\lambda}{2} \right) &\leq \ln \frac{\bar{k}}{\hat{k}_2} \\ -\alpha_2 \left(\frac{\lambda}{2} + \overline{\Lambda}_o \right) &\geq \ln \frac{1}{2} \end{aligned}$$

Since $\frac{\bar{k}}{\hat{k}_2} > 1$, this is always feasible. Thus, there is a fixed point, and we are done. ■

References

Acemoglu, Daron, 2001, Good jobs versus bad jobs, *Journal of Labor Economics* 19, 1–22.

Acemoglu, Daron, and Robert Shimer, 2000, Wage and technology dispersion, *The Review of Economic Studies* 67, 585–607.

Akerlof, George A. and Lawrence F. Katz, 1989, Workers' trust funds and the logic of wage profiles, *The Quarterly Journal of Economics* 104, 525–536.

Baranchuk, Nina, Glenn MacDonald, and Jun Yang, 2008, The economics of super managers, working paper, Indiana University.

Bulow, Jeremy I., and Lawrence H. Summers, 1986, A theory of dual labor markets with application to industrial policy, discrimination, and Keynesian unemployment, *Journal of Labor Economics* 4, 376–414.

Edmans, Alex, Xavier Gabaix, and Augustin Landier, forthcoming, A multiplicative model of optimal CEO incentives in market equilibrium, *Review of Financial Studies.*

Gayle , George-Levi, and Robert A. Miller, forthcoming, Has moral hazard become a more important factor in managerial compensation?, *American Economic Review.*

Hutchens, Robert, 1986, Delayed payment contracts and a firm's propensity to hire older workers, *Journal of Labor Economics* 4, 439–457.

Krueger, Alan B., and Lawrence H. Summers, 1988, Efficiency wages and the inter-industry wage structure, *Econometrica* 56, 259–293.

Lazear, Edward P., 1981, Agency, earnings profiles, productivity, and hours restrictions, *The American Economic Review* 71, 606–620.

MacLeod, W. Bentley, and James M. Malcomson (1998), Motivation and markets, *American Economic Review* 88, 388–411.

Medoff , James L., and Katharine G. Abraham, 1980, Experience, performance, and earnings, *The Quarterly Journal of Economics* 95, 703–736.

Moen, Espen R., and Åsa Rosén, 2007, Incentives in competitive search equilibrium, working paper, Stockholm University.

Moen, Espen R., and Åsa Rosén, 2006, Deferred compensation and turnover, working paper, Stockholm University.

Moen, Espen R., and Åsa Rosén, forthcoming, Equilibrium incentive contracts and efficiency wages, *Journal of European Economic Association.*

Oyer, Paul, forthcoming, The making of an investment banker: Stock market shocks, career choice, and life-time income, *Journal of Finance.*

Sannikov, Yuliy, forthcoming, A continuous-time version of the principal-agent problem, *Review of Economic Studies.*

Shapiro, Carl, and Joseph E. Stiglitz, Equilibrium unemployment as a worker discipline device, *The American Economic Review* 74, 433–444.

Walsh, Frank , 1999, A multisector model of efficiency wages, *Journal of Labor Economics* 17, 351–376.