Retained Earnings Dynamic, Internal Promotions and Walrasian Equilibrium

Pablo F. Beker

No 813

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Retained Earnings Dynamic, Internal Promotions and Walrasian Equilibrium

Pablo F. Beker★

Department of Economics
University of Warwick

August, 2007.

Correspondence to:

Pablo F. Beker
Department of Economics
University of Warwick
Coventry CV4 7AL
United Kingdom
Email: P.Beker@warwick.ac.uk
Tel: +44 (0) 24 761 50588
Fax: +44 (0) 24 765 23032.

★ This paper is a revised version of chapter III of my Ph.D. dissertation submitted to the Graduate School of Cornell University in January 2002 and supervised by Prof. David A. Easley. I thank him for comments on a previous version of this paper. I am also grateful to Subir Chattopadhyay, for many helpful discussions and suggestions that greatly improved the paper, and to an Associate Editor, an anonymous referee and Victor A. Beker for useful comments. I also benefited from discussions with the participants of the First Brazilian Workshop of the Game Theory Society held in Sao Paulo, the XIII World Congress of the IEA held in Lisbon, the European Economic Association 2004 Annual Congress in Madrid, and the XIX Simposio de Análisis Económico at Unversidad de Navarra and seminars at the Universities of Valencia, Alicante, El Colegio de Méjico and Carlos III. All remaining errors are mine. Support by the Spanish Ministry of Science and Technology, Grant N^o BEC2001-0980, as well as from the Ivie is gratefully acknowledged.

Abstract

In the early stages of the process of industry evolution, firms are financially constrained and might pay different wages to workers according to their expectations about the prospects for advancement offered by each firm's job ladder. This paper argues that, nevertheless, if the output market is competitive, the *positive predictions* of the perfectly competitive model are still a good description of the long run outcome. If firms maximize the discounted sum of constrained profits, financing expenditure out of retained earnings, profits are driven down to zero as the perfectly competitive model predicts. *Ex ante* identical firms may follow different growth paths in which workers work for a lower entry-wage in firms expected to grow more. In the steady state, however, workers performing the same job, in *ex-ante* identical firms, receive the same wage. I explain when the long run outcome is efficient, when it is not, and why firms that produce inefficiently might drive the efficient ones out of the market even when the steady state has the positive properties of a Walrasian equilibrium. To some extent, it is not technological efficiency but workers' self-fulfilling expectations about their prospects for advancement within the firm that explains which firms have lower unit costs, grow more, and dominate the market.

Key words: Industry Evolution - Market Selection Hypothesis - Production under Incomplete Markets - Retained Earnings Dynamic - Self-Fulfilling Expectations - Internal Labor Markets

J.E.L. Classification Numbers: D21, D52, D61, D84, D92, J41

1. INTRODUCTION

Economic theory predicts that in a market where many firms sell an homogeneous product, at least in the long run, each firm produces the quantity that maximizes profits and profits vanish. Although most economists agree with this description of the long run outcome of the process of industry evolution, much less consensus has been achieve on what forces lead an industry to that state. The theory of industry equilibrium in competitive markets relies on the existence of perfect markets for inputs and credit to explain why profits are dissipated. Indeed, if there is a complete set of perfectly competitive financial markets, each firm maximizes its market value, the markets for inputs are perfectly competitive, there are no turnover costs and there is either free entry or the technology displays constant returns to scale then profits are zero from the start and each firm produces at the profit maximizing level.

In sharp contrast with these assumptions, however, the empirical evidence suggests that new firms are financially constrained and the labor market, rather than being in a Walrasian equilibrium from the start, is better characterized by social institutions not considered in the theory of the firm under perfect competition.

Indeed, on the one hand, the problems of asymmetric information identified by authors like Stiglitz and Weiss [13] as the main explanation for the failure of the credit market are important in the early stages of the process of industry evolution. Consequently, financing through retained earnings is the norm rather than the exception.[1] The lack of access to credit prevents firms from achieving their optimal size from the start and explains why it takes time for profits to be dissipated. On the other hand, workers tend to be attached to the same firm for long periods, firms carry out most of the training of their employees and prefer to promote employees rather than recruiting new workers. Using the term made popular by Doeringer and Piore [8], firms set up an internal labor market.

It is apparent that modern industries display many features that are not taken into account in the static model but are key to understanding why industry evolution takes time and how wages evolve. Therefore, the standard description of firm and industry behavior is at best the description of a steady state of some growth dynamics. Economists like Alchian [1] and Friedman [10] recognized this a long time ago. Nelson and Winter [12, and references therein] were the first to provide a formal explanation of how such a steady state can be attained even if no firm follows a profit maximization rule. The key assumption in their work is that firms that make positive profits expand, those that make zero profits do not change capacity while those that make losses contract and search for new decision rules, a dynamic that can be motivated by the use of retained earnings to finance investment. However, Blume and Easley [7] show that even though such retained earnings dynamic may explain why firms that do not maximize profits are driven out, it may not converge to a Walrasian equilibrium. In their model, capitalists finance production out of retained earnings, so that firms with uniformly smaller profits are driven out, and the interior steady states of the retained earnings dynamic are competitive equilibria and are Pareto optimal. They provide examples in which firms produce two goods operating different technologies and they show that if the

[1] To quote Allen and Gale [3]: “Perhaps the most striking point [...] is that in all countries [US, UK, France and Germany] except Japan, retained earnings are the most important source of funds. External finance is simply not that important” (p. 76)

goods are sufficiently complementary, then the steady state of the retained earnings dynamic may be unstable.[2]

The work of Nelson and Winter and Blume and Easley, however, focuses on the role of the retained earnings dynamic as a substitute for market completeness when the labor market is in a Walrasian equilibrium from the start. In many industries, instead, the existence of training costs and firm specific abilities lead firms to set up internal labor markets where wages exceed those of competing industries. This is typically the case for skill intensive jobs at the top of the progression line. Since workers anticipate they may progress through the promotion line and obtain those high wages in the future, intuitively, the better the prospects for advancement displayed by the firm, *ceteris paribus*, the lower the worker's reservation entry-wage is. This insight introduces a potentially self-fulfilling aspect in the process of industry evolution. Indeed, since firms rely on internal funds, firms believed to have better growth potential pay, *ceteris paribus*, lower wages, and have more revenue, end up promoting more workers and producing more, fulfilling workers' expectations. If *ex-ante* identical firms follow different growth paths, does the industry converge to a steady state? What are the efficiency properties of the steady state? Which firms pay lower wages? Is there an unambiguous positive relationship between technological efficiency and growth rates? These are some of the questions addressed in this work.

This paper argues that when firms finance expenditure out of retained earnings and the internal labor market arises as a cost minimizing institution (due to firm specific abilities and costly training), the industry converges to a steady state that is Walrasian-like in the sense that profits are dissipated, firms do not face financial constraints and markets clear. However, this steady state need not efficient. Indeed, along the transition, inefficient firms may display better growth prospects, pay lower entry-wages and grow more than the efficient ones. Adjustment costs do no play any role in my analysis because firms do not face a shortage in the internal supply of skilled workers along the process of industry evolution. Instead, I focus on the role of workers' expectations in shaping factor prices, an aspect that has not been addressed yet in the literature of industry evolution towards a Walrasian equilibrium.

I use a partial equilibrium model of industry evolution with long live firms that operate a two-task technology with constant returns and infinitely many overlapping generations of workers who live for two periods and consume out of wages. A worker who performs the first task when young develops high ability with positive probability. Only high ability workers who undergo training can perform the second task when old. Training is costly for the firm but is costless for the worker who is free to move after the training process has ended. A worker who does not switch firms, is more productive in the second task than a worker trained by another firm. A key ingredient is that there are always firms that are yet to train when others have ended their training process and so the former can save the training cost by hiring workers trained by the latter. Firms compete for skilled workers a la Bertrand and so the

[2] The intuition behind this result is as follows. Suppose firm 1's technology displays constant returns to scale and that with 1 unit of input at t it produces more of good 1 and less of good 2 at $t+1$ than firm 2. When firm 1's financial capital exceeds its steady state level, total output of good 1 increases and total output of good 2 decreases. If the goods are sufficiently complementary, the extra output of good 1 and the corresponding reduction of good 2 reduces the market clearing price of good 1 so much that firm 1 suffers a large loss. In particular, firm 1's retained earnings can fall below its steady state level, while the opposite holds for firm 2 which mainly produces good 2. This causes the opposite response in the next period, with firm 2 producing more of good 2 than in the steady state. When the goods are sufficiently complementary, this cycle of profits and loses produces cycles in the levels of financial capital that do not damp out and allocative efficiency fails. They also have an example with four firms where the equilibrium does not converge to a steady state and technological efficiency fails.

skilled worker's wage is his best outside offer. The higher the training cost is or the more general the skill is, the higher is the firms' willingness to pay for an externally trained worker and, therefore, the higher is his wage upon promotion. When the latter exceeds his wage at home, his reservation entry-wage depends on the worker's beliefs about the firm's promotion rate. Firms take output prices as given, cannot borrow and allocate their assets either to finance production or to a risk free activity. As long as profits are positive, firms find optimal to fully allocate their assets to finance production, driving prices down and pushing profits to the steady state level.

I show, by means of an example, that *ex-ante* identical firms can follow different growth paths towards the steady state. *Ceteris paribus*, firms expected to grow faster hire workers at a lower entry-wage and so technological efficiency fails along the transition. However, technological efficiency does hold in the steady state since growth rates are identical across firms. In the steady state, therefore, not only firms are not financially constrained and make zero profits but also workers who carry out identical jobs receive the same wage regardless of the firm that hires them and the labor markets clears *as if* it were in a Walrasian equilibrium. Allocative efficiency, instead, holds in the steady state if and only if the training cost is so low or the ability so firm specific that the skilled worker's wage equals his wage at home and so entry-wages are identical across productive activities. Otherwise, too little is produced compared to the efficient allocation of resources. The failure of technological and allocative efficiency is due both to the absence of a perfect credit market as well as the impossibility of paying old workers their opportunity cost out of the industry.

I also consider the case of firms with different technologies. Although economists long time ago recognized that firms with lower costs grow more, it is usually assumed that cost differentials stem from technological factors. A reason that has been overlooked is that, *ceteris paribus*, those firms believed to display better growth prospects can hire workers at a lower wage. Since the workers' willingness to work for a lower entry-wage can more than compensate for the disadvantage introduced by an inefficient technology, even firms that produce inefficiently may end up dominating a market if workers believe they display sufficiently better prospects than the efficient ones. Can it happen in an equilibrium converging to a Walrasian-like state? I construct an example in which profits vanish, worker's expectations are fulfilled and, nevertheless, inefficient firms grow more and dominate the market in terms of market share. If at the early stages of the process of industry evolution workers are optimistic enough about prospects for advancement offered by the firms that produce inefficiently, *almost* all workers end up employed by inefficient firms in the long run. Therefore, firms do not face financial constraints and make zero profits, markets clear and *almost* all workers performing the same job receive the same wage, as in a Walrasian equilibrium.

In section 2 I define the game of imperfect information played by the firms and the infinite generations of workers for a fixed sequence of output prices. I characterize the equilibrium path of its Perfect Bayesian Equilibrium (PBE) and prove existence in sections 3 and 4, respectively. In section 5, I define an Industry Equilibrium (IE) as a PBE where output prices clear the market. In section 6 and 7, I show there is an IE in which *ex-ante* identical firms follow identical growth paths and I discuss the efficiency properties of the IE for the case of *ex-ante* identical and heterogeneous firms, respectively. Proofs are in the Appendix.

2. THE MODEL

At date zero, the industry adopts a new technology based in two tasks that use only labor as input. If q_1 and q_2 denote the level at which the two tasks are performed, then the output level, q, is given by

$$q = q_1^{\alpha} \cdot q_2^{1-\alpha} \qquad \text{where } \; 0 < \alpha < 1.$$

Task 1 requires a skill that is not industry specific. If l is the number of workers employed in task 1, then $q_1(l) = l$.[3] Every worker develops a new ability while performing the first task. Ability is a random variable that takes only two values: high or low; ability turns out to be high with probability $\lambda \in (0, 1)$. Only high ability workers who have received training develop an industry specific skill necessary to perform the second task. A firm can find a worker who is suitable to perform the second task at $t + 1$ from one of two sources:[4]

1. The employees that performed task 1 in the firm at t and developed high ability.[5] They can be trained at the beginning of $t + 1$, at a unit cost of c, to perform task 2 during $t + 1$. If the firm hires them, it is said to promote workers internally. If every worker performing task 2 has been promoted internally, the firm is said to have a *closed internal labor market with one entry port.*

2. The employees that performed task 1 in other firms in the industry at t. If one of these workers decides to move after being trained, he does not need additional training to perform the second task in his new job. However, he is not as productive as a skilled employee who worked in the same firm when young: e skilled workers that switch firms are as productive as $\frac{e}{1+\theta}$, with $\theta > 0$, skilled employees who are promoted internally. Larger values of θ correspond to greater firm specificity of the skill. A firm that employs them is said to hire workers externally.

If s and e are the number of internally promoted and externally hired workers, respectively, then $q_2(s, e) = s + \frac{e}{1+\theta}$ denotes the level of activity of the second task. A firm that employs (l, s, e) workers at date t, obtains $q(l, s, e; \alpha) \equiv q_1(l)^{\alpha} \cdot q_2(s, e)^{1-\alpha}$ units of output at $t + 1$.

2.1 Workers

Every period $t \geq 0$, a new generation of workers, who live for two periods, enters the labor force. Workers do not consume the good produced by this industry. They only face uncertainty about their ability and, therefore, about their wage when old. Workers have preferences over random bundles of the numeraire that display risk neutrality and discount the future at rate $0 < \beta < 1$. A worker who does not work in this industry can work at home with expected lifetime utility $\overline{w}_1 + \beta \cdot \overline{w}_2$, when young, and $\overline{w}_2 \geq \overline{w}_1 > 0$, when old. Without loss of generality, one may think that $\overline{w}_1$ and $\overline{w}_2$ are the productivities of a young and an old worker in the production of the numeraire. Since workers cannot borrow, they consume out of wages. I assume $\frac{c}{\theta}$ is large relatively to $\overline{w}_1$.

Assumption AW: $\overline{w}_1, \overline{w}_2, \beta$, and $\frac{c}{\theta}$ are such that $\overline{w}_2 < \frac{c}{\theta}$ and $\overline{w}_1 \leq \lambda \cdot \beta \cdot \left[\frac{c}{\theta} - \overline{w}_2\right]$.

[3] It would be more appropriate to say that l is the measure of workers hired by the firm. The same applies to all other types of labor.

[4] There is a third possibility. A firm could screen a worker who did the first task in other industry when young, to learn whether he has high ability or not. I assume screening costs are prohibitively high. See Doeringer and Piore [8, p. 31] for arguments supporting this assumption.

[5] Doeringer and Piore emphasize this point [8, p. 31].

2.2 Efficient Allocations

The demand for the good, $D(p)$, has standard properties.

Assumption AD: $D : \Re_+ \rightarrow \Re_+$ is continuous and strictly decreasing for all p such that $D(p) > 0$, $\lim_{p\rightarrow\infty} D(p) = 0$ and $D(p) = 0 \Rightarrow p \cdot \left(\frac{\alpha}{\overline{w}_1}\right)^{\alpha} \cdot \left(\frac{1-\alpha}{\max\{\overline{w}_2, c/\theta\}+c}\right)^{1-\alpha} > r$.[6]
where$1 < r < \frac{1}{\beta}$ is the gross rate of return on a risk free investment opportunity.[7]

Since this is a partial equilibrium model, to make efficiency judgments one has to make some additional assumptions. I assume that consumer surplus is an adequate measure of welfare and $\frac{1}{r}$ is the socially optimal discount rate. Since $\overline{w}_1$ and $\overline{w}_2$ are the productivities of young and old workers out of this industry, they measure the social cost of allocating workers to this industry. At any $t \geq 0$, there are only two relevant types of labor for the planner: the young workers who perform task 1 and the old workers who performed task 1 in this industry when young. An industry is technologically efficient if more output cannot be produced using the same amount of every input and strictly less of one of them. As usual, the set of efficient allocations is the solution to the following Social Planner's problem where $CS(q) \equiv \int_0^q D^{-1}(x)\,dx$ is the Marshallian Consumer Surplus,[8]

$$\max_{l_t, s_t \geq 0} \sum_{t=0}^{\infty} \left(\tfrac{1}{r}\right)^t \cdot \left[\tfrac{1}{r} \cdot CS\left(l_t^{\alpha} \cdot s_t^{1-\alpha}\right) - \overline{w}_1 \cdot l_t - (\overline{w}_2 + c) \cdot s_t\right] \quad s.t. \quad s_t \leq \lambda \cdot l_{t-1} \tag{1}$$

Let $p^*(\alpha) \equiv r \cdot \left(\frac{\overline{w}_1}{\alpha}\right)^{\alpha} \cdot \left(\frac{\overline{w}_2+c}{1-\alpha}\right)^{1-\alpha}$ and $Q^*(\alpha) = D(p^*(\alpha))$. Lemma 2.1 characterizes the set of efficient allocations for those parameters such that in the solution to (1) the constraint does not bind.[9]

Lemma 2.1 *If $\alpha > \frac{\overline{w}_1}{\overline{w}_1+\lambda\cdot(\overline{w}_2+c)}$ then $Q^*(\alpha)$ is the allocative efficient level of output while the allocative efficient levels of labor are $l_t = \left(\frac{\alpha}{1-\alpha} \cdot \frac{\overline{w}_2+c}{\overline{w}_1}\right)^{1-\alpha} \cdot Q^*(\alpha)$ and $s_t = \left(\frac{1-\alpha}{\alpha} \cdot \frac{\overline{w}_1}{\overline{w}_2+c}\right)^{\alpha} \cdot Q^*(\alpha)$.*

2.3 Firms

Firms have names in the set $\mathcal{I}$ of rational numbers in $(0,1)$, take the output price sequence $P = \{p_t\}_{t=0}^{\infty} \in \Re_+^{\infty}$ as given, and are endowed with $a_0 > 0$ units of the numeraire and $\lambda \cdot l_{-1} \geq \frac{1-\alpha}{\overline{w}_2+c} \cdot a_0$ trainees.[10] The distribution of workers across firms satisfies a law of large numbers at each date: if firm i employs l workers in task 1 at date t, a fraction λ of these workers develops high ability.[11] Firms cannot borrow. At every $t \geq 0$, each firm collects earnings and decides how much of its assets to invest in an alternative activity with gross rate of return $r > 1$ and what part to allocate as financial capital to hire inputs.[12] Figure 1 illustrates the timing of decisions.

[6] The last condition implies there are prices so that firms that pay wages $\overline{w}_1$ and $\max\left\{\overline{w}_2, \frac{c}{\theta}\right\}$ to workers who perform the first and second task, respectively, can make positive profits. This assumption will guarantee that equilibrium output is not zero.

[7] To simplify the analysis I do not consider the case in which $\beta = \frac{1}{r}$ even though all the results in this paper extend to that case.

[8] A standard justification is that agents who consume the good produced in this industry have quasilinear preferences.

[9] This is the appropriate benchmark because in all the equilibria I analyze later, the constraint does not bind either.

[10] One can think that firms have been operating for a while, perhaps using another technology based only in task 1, and know the ability of those workers it employed before.

[11] Since independence has no role in this model, the argument in Feldman and Gilles [9] implies that there exists a distribution of workers for which the law of large numbers holds in every Borel set.

[12] For the rest of the paper, I take this alternative activity as lending at the interest rate r.

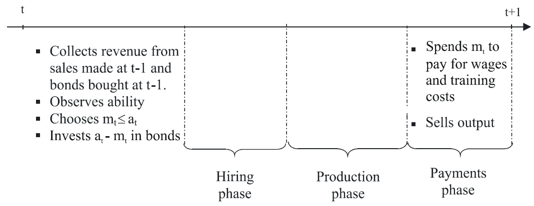

Figure 1. Timing of decisions

2.4 The hiring process

In principle, there is a large set of labor contracts a firm could offer to its workers. For example, one could imagine a contract in which the firm assigns a young worker to task 1, pays him a certain wage at date t and promises future wages contingent on being promoted or not. One could even think of a contract where the firm details the promotion rate at $t+1$, as in Malcomson [11]. However, many contracts like these are not implementable because of a lack of commitment by the workers or the firm. Therefore, I restrict the analysis to spot contracts.

Assumption AC: Firms can neither commit to a wage in the event a worker is promoted nor to a promotion rate.

At date $t \geq 0$ every firm faces an infinite supply of young workers each of whom is contacted by just one firm. If training were completely firm specific, one would expect the wage of a promoted worker to be $\overline{w}_2$ because no other firm in this industry would be willing to hire such worker. In this model, however, the second task can be performed by workers hired internally or externally. Although internally promoted workers are more productive than the externally trained ones, a firm who hires externally avoids the training cost c. Evidently, for trained workers to have an outside opportunity within the industry it is necessary that there be a firm that is yet to train when his training process has ended. This seems very realistic. To capture this idea, I assume that between dates t and $t+1$ firms carry out the training process sequentially. $\mathcal{S}_i$ and $\mathcal{D}_i$ denote the set of firms that train workers right before and after firm i does it. Of course, $\mathcal{S}_i \cap \mathcal{D}_i = \varnothing$. For simplicity, I also assume $\#\mathcal{S}_i = \#\mathcal{D}_i = N$.

The interaction between the firms and the successive generations of workers define an extensive game $\Gamma(P)$. At every date $t \geq 0$, the game $\Gamma(P)$ consists of three stages:

■ $1^{\underline{st}}$ stage: Firms, simultaneously, decide the number of vacancies they open for the first task ($l_{i,t}$) and externally trained workers ($e_{i,t}$), the wage offers associated with the first task ($w_{i,t}$), internal promotions ($v_{i,t}$) and externally trained workers ($v^e_{i,t}$), the financial capital ($m_{i,t}$) and the bond holdings ($b_{i,t}$). That is, every firm $i \in \mathcal{I}$ chooses a vector $d_{i,t} = \left(l_{i,t}, e_{i,t}, w_{i,t}, v_{i,t}, v^e_{i,t}, m_{i,t}, b_{i,t}\right) \in \Re^7_+$ such that

$$w_{i,t} \cdot l_{i,t} + v^e_{i,t} \cdot e_{i,t} = m_{i,t} \tag{2}$$

Condition (2) limits the vacancies for young workers and externally trained workers to what firm i can finance at the announced wages using all its financial capital. In addition, firm i's financial capital and bond holdings must be non-negative and add up to the firm's assets ($a_{i,t}$),

$$(m_{i,t}, b_{i,t}) \in \Re^2_+ \tag{3}$$

$$m_{i,t} + b_{i,t} = a_{i,t} \tag{4}$$

■ $2^{\underline{nd}}$ stage: Young workers, simultaneously, decide whether to accept (A) or reject (R) employment.

<u>ALLOCATION OF 2ND TASK JOB OFFERS</u>: Firm i's skilled workers get offers only from firms in $\mathcal{D}_i$ and firm i offers are received only by skilled workers of firms in $\mathcal{S}_i$. A worker receives at most one offer from each firm. If firm $j \in D_i$ announces $e_{j,t}$ vacancies for externally trained workers, at most that number of workers trained by firms in $\mathcal{S}_j$ gets an offer. Let $\widehat{v}_{j,t}^e$ be a random variable taking values in $\left\{v_{j,t}^e, 0\right\}$. Each worker trained by firm i observes a realization of $\widehat{v}_{j,t}^e$ for each $j \in \mathcal{D}_i$ with the interpretation that those who observe $v_{j,t}^e$ are being offered a job in firm j at wage $v_{j,t}^e$ and those who observe 0 not.

Let $s_{i,t-1}$ be the number of workers that were offered training at $t-1$ and $\widehat{l}_{i,t-1}, \widehat{s}_{i,t-1}$ and $\widehat{e}_{i,t-1}$ be the number who accepted employment at $t-1$ in firm i. Then,[13]

$$\left(\widehat{l}_{i,t-1}, \widehat{s}_{i,t-1}, \widehat{e}_{i,t-1}\right) \leq \left(l_{i,t-1}, s_{i,t-1}, e_{i,t-1}\right), \tag{5}$$

where $\widehat{l}_{i,-1} = l_{-1}$ and $\widehat{s}_{i,-1} = \widehat{e}_{i,-1} = 0$. Because the number of workers who are offered training cannot exceed the number of high ability workers, the firm faces the following "internal labor market" constraint at t:

$$0 \leq s_{i,t} \leq \lambda \cdot \widehat{l}_{i,t-1}. \tag{6}$$

If $s_{i,t} < \lambda \cdot \widehat{l}_{i,t-1}$, the firm decides at random who is offered training because, from its point of view, high ability workers are homogeneous. Since internal promotions must be financed with the money left after financing externally trained workers who accept employment in firm i, $\widehat{e}_{i,t}$, then[14]

$$s_{i,t} = \begin{cases} \min\left\{\frac{m_{i,t} - w_{i,t} \cdot l_{i,t} - v_{i,t}^e \cdot \widehat{e}_{i,t}}{v_{i,t} + c}, \lambda \cdot \widehat{l}_{i,t-1}\right\} & \text{if } q\left(\widehat{l}_{i,t-1}, \widehat{s}_{i,t-1}, \widehat{e}_{i,t-1}, \alpha_i\right) > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

where the second line re ects that young workers who accepted employment in a firm that produced zero at $t-1$ did not develop any ability because they did not perform the first task.

■ $3^{\underline{rd}}$ stage: Each old worker trained by firm i observes a random draw of $\widehat{v}_{j,t}^e$ for each $j \in \mathcal{D}_i$ and decides whether to stay (choose i), move to some firm $j \in \mathcal{D}_i$ that made him an offer or work at home (choose o).

Finally, firm i's assets at date $t+1$ are

$$a_{i,t+1} \equiv p_t \cdot q\left(\widehat{l}_{i,t}, \widehat{s}_{i,t}, \widehat{e}_{i,t}; \alpha_i\right) + \left[m_{i,t} - c \cdot s_{i,t} - w_{i,t} \cdot \widehat{l}_{i,t} - v_{i,t} \cdot \widehat{s}_{i,t} - v_{i,t}^e \cdot \widehat{e}_{i,t}\right] + r \cdot b_{i,t} \tag{8}$$

Implicit in condition (8) is that even though the financial capital the firm would use to pay the wage of those workers that reject employment cannot be invested in bonds, it can be stored and spent the following period.

[13] For any $x, y \in \Re_+^n$, $x \geq y$ if $x_k \geq y_k$ for $k = 1, ..., n$.

[14] Since training is costless for the worker, I assume every worker who is offered training accepts and, therefore, I use $s_{i,t}$ to denote also the number of workers who accept training.

2.4.1 Formal description of the extensive form game $\Gamma(P)$.

For each $t \geq 0$, $i \in \mathcal{I}$, $d_{i,t} \in \Re^7_+$ and $\left(s_{i,t}, \widehat{l}_{i,t}, \widehat{s}_{i,t}, \widehat{e}_{i,t}\right) \in \Re^4_+$, define $h_{i,t} \equiv \left(d_{i,t}, s_{i,t}, \widehat{l}_{i,t}, \widehat{s}_{i,t}, \widehat{e}_{i,t}\right)$. H_i^∞ is the set of sequences $(h_{i,0}, h_{i,1}, ..., h_{i,t}, ...)$ such that (2) - (8) holds at every date $t \geq 0$ and $H^\infty \equiv \underset{i\in\mathcal{I}}{\times} H_i^\infty$ is the set of play paths. A typical element of H^∞ is a sequence $h = (h_0, h_1, ..., h_t, ...)$ where $h_t = \{h_{i,t}\}_{i\in\mathcal{I}}$ for every $t \geq 0$. For each $h \in H^\infty$, $h_i^t(h) \equiv (h_{i,0}, h_{i,1}, ..., h_{i,t-1})$ denotes the partial history of firm i up to date t, $h^t(h) \equiv (h_0, h_1, ..., h_{t-1})$ denotes the partial history of the game up to date t, $\overline{h_i^t}(h) \equiv \left(h_{i,0}, h_{i,1}, ..., h_{i,t-1}, d'_{i,t}\right)$ is firm i's partial history up to date t including firm i's actions at t, and $\overline{h^t}(h) \equiv (h_0, h_1, ..., h_{t-1}, d'_t)$, where $d'_t = \left(d'_{1,t}, ..., d'_{i,t}, ...\right)$, is the partial history of the game up to date t including every firm's actions at t. H_i^t is the set of partial histories of firm i up to date t, H^t is the set of partial histories up to date t and $\overline{H^t}$ is the set of partial histories up to date t including the actions of firms at t.

At date $t \geq 0$, each firm recalls its own past actions and observes the number of workers who accepted employment in that firm in the past. Therefore, the information set of firm i after observing $x_i^t \in H_i^t$ is

$$I_{i,t}\left(x_i^t\right) \equiv \left\{x^t \in H^t : \exists h \in H^\infty,\, h^t(h) = x^t \text{ and } h_i^t(h) = x_i^t\right\}.$$

The initial assets of firm i are $a_{i,0}(h) = a_0$ and for every $t \geq 0$, firm i's assets at $t+1$ are

$$a_{i,t+1}(h) \equiv p_t \cdot q_{i,t}(h) + \left[m_{i,t}(h) - c \cdot s_{i,t}(h) - w_{i,t}(h) \cdot \widehat{l}_{i,t}(h) - v_{i,t}(h) \cdot \widehat{s}_{i,t}(h) - v^e_{i,t}(h) \cdot \widehat{e}_{i,t}(h)\right] + r \cdot b_{i,t}(h), \tag{9}$$

where $q_{i,t}(h) \equiv q\left(\widehat{l}_{i,t}(h), \widehat{s}_{i,t}(h), \widehat{e}_{i,t}(h); \alpha_i\right)$ and $R_{i,t}(h) = \frac{a_{i,t+1}(h)}{a_{i,t}(h)}$ is firm i's production level and rate of return at date t on path h, respectively. The set of actions available to firm i on path h at date t is

$$\mathcal{A}_{i,t}(h) = \left\{(l, e, w, v, v^e, m, b) \in \Re^7_+ : w \cdot l + v^e \cdot e = m,\, m + b = a_{i,t}(h)\right\}.$$

For $x_i^t \in H_i^t$, let $\mathcal{A}_{i,t}\left(x_i^t\right) = \mathcal{A}_{i,t}(h)$ if $h_i^t(h) = x_i^t$. $\mathcal{A}_{i,t} = \cup_{h\in H^\infty} \mathcal{A}_{i,t}(h)$ is the set of actions of firm i at t.

Since a young worker contacted by firm i at date t only observes the wage offer $w \geq 0$, his information set is

$$I^1_{i,t}(w) \equiv \left\{\overline{x^t} \in \overline{H^t} : \exists h \in H^\infty \text{ such that } \overline{h^t}(h) = \overline{x^t} \text{ and } w_{i,t}(h) = w\right\}.$$

The function $\chi_{i,t}(h)$ indicates whether the number of trainees in firm i at date t on path h is limited by its assets ($\chi_{i,t}(h) = 0$) or by its internal supply of high ability workers ($\chi_{i,t}(h) = 1$).[15] Let $\chi_t^i(h)$ be the vector in $\Re^N_+$ with coordinates $\chi^i_{j_n,t}(h)$, $j_n \in \mathcal{D}_i$. For fixed $t \geq 0$, $h \in H^\infty$ and $i \in \mathcal{I}$, $\widehat{v}_t^i(h)$ is the random vector in $\Re^N_+$ with coordinates $\widehat{v}^e_{j_n,t}(h)$, $j_n \in \mathcal{D}_i$, and probability distribution $\eta_{i,t}(h)$. Each worker trained by firm i at date t observes $\left(v_{i,t}(h), \chi_t^i(h)\right) \in \Re^{N+1}_+$ and a random draw of $\widehat{v}_t^i(h)$. So, he observes a realization of $\overline{v}_{i,t}(h) \equiv \left[v_{i,t}(h), \widehat{v}_t^i(h), \chi_t^i(h)\right]$.

[15] That is, $\chi_{i,t}(h) = 0$ if $\frac{m_{i,t}(h) - w_{i,t}(h) \cdot l_{i,t}(h)}{v_{i,t}(h) + c} < \lambda \cdot \widehat{l}_{i,t-1}(h)$ and $\chi_{i,t}(h) = 1$ otherwise.

2.4.2 Strategies

Firm i's strategy is a sequence $f_i \equiv \{f_{i,t}\}_{t=0}^{\infty}$ where $f_{i,t} : H_i^t \to \mathcal{A}_{i,t}$ satisfies $f_{i,t}\left(x_i^t\right) \in \mathcal{A}_{i,t}\left(x_i^t\right)$ and $\mathbb{F}_i$ is firm i's set of pure strategies.

The strategy of a worker born at t who is contacted by firm i is a pair $\sigma_{i,t} = \left(\sigma_{i,t}^1, \sigma_{i,t+1}^2\right)$ where $\sigma_{i,t}^1 : \Re_+ \to \{A, R\}$ is his response when young after being offered w by firm i and $\sigma_{i,t+1}^2 : \Re_+^{2 \cdot N+1} \to \mathcal{D}_i \cup \{i, o\}$ is his response when old after being trained by firm i and observing $\overline{v}_{i,t}(h)$. $\mathbb{W}_{i,t}$ is the set of strategies of the workers born at t. $\sigma_i \equiv \left\{\left(\sigma_{i,t}^1, \sigma_{i,t}^2\right)\right\}_{t=0}^{\infty} \in \mathbb{W}_i \equiv \times_{t=0}^{\infty} \mathbb{W}_{i,t}$ is the collection of strategies the infinite generations of workers play against firm i when all workers of a given generation play the same strategy against firm i.

A profile of strategies γ is a collection $\{f_i, \sigma_i\}_{i \in \mathcal{I}}$ such that $(f_i, \sigma_i) \in \mathbb{F}_i \times \mathbb{W}_i$ for every $i \in \mathcal{I}$. γ_{-i} and γ'_{-i} are the collections $\{\sigma_i, f_j, \sigma_j\}_{j \in \mathcal{I}, j \neq i}$ and $\{f_i, f_j, \sigma_j\}_{j \in \mathcal{I}, j \neq i}$, respectively. For a fixed profile γ, $h^*(\gamma) \in H^\infty$ and $h^*(x, \gamma) \in H^\infty$ denote the path of play and the path of play after partial history $x \in H^t \cup \overline{H^t}$, respectively.

2.4.3 Payoffs, beliefs and equilibrium

At date t, an old worker, who underwent training in firm i and observes a realization of $\overline{v}_{i,t}(h)$, can stay in firm i and obtain wage $v_{i,t}$, or join firm $j \in \mathcal{D}_i$ and receive wage $\widehat{v}_{j,t}^e$, or work at home and get $\overline{w}_2$. His payoff is

$$\widehat{u}_{i,t}\left(\sigma_{i,t}^2\right)\left[\overline{v}_{i,t}(h)\right] \equiv \begin{cases} v_{i,t}(h) & \text{if } \sigma_{i,t}^2\left(\overline{v}_{i,t}(h)\right) = i. \\ \widehat{v}_{j,t}^e(h) & \text{if } \sigma_{i,t}^2\left(\overline{v}_{i,t}(h)\right) = j \in \mathcal{D}_i. \\ \overline{w}_2 & \text{otherwise.} \end{cases}$$

which is independent of both the history of the game and the other players' strategies. The expected payoff of an old worker at date t on path h before observing $\overline{v}_{i,t}(h)$ is $u_{i,t}\left(\sigma_{i,t}^2, h\right) \equiv E_{\eta_{i,t}(h)}\left[\widehat{u}_{i,t}\left(\sigma_{i,t}^2\right)\left[\overline{v}_{i,t}(h)\right]\right]$.

The payoff of firms and young workers, instead, does depend on the history up to that date. Therefore, one needs to specify their beliefs about the history of the game conditional on their information up to that date.

Let I_t and I_t^1 be the sets of all information sets of firms and young workers at date t, respectively. I is the set of all information sets of firms and young workers. Since each $\xi \in I$ is a countable product of rectangles in the Euclidean space, in order to define beliefs one needs to define an appropriate measurable structure. Let $\mathcal{F}(\xi)$ be the σ-algebra generated by the elements of ξ and $(\xi, \mathcal{F}(\xi))$ be a measurable space. A system of beliefs is a collection $\Psi \equiv \{\Psi_\xi,\ \xi \in I\}$ such that for every $\xi \in I$, $\Psi_\xi : \mathcal{F}(\xi) \mapsto [0,1]$ is a probability measure on $(\xi, \mathcal{F}(\xi))$.

Definition 2.1 *A system of beliefs Ψ is consistent with a profile of strategies γ if the following conditions hold*
(i) if $\xi = I_{i,t}\left(h_i^t(h^*)\right)$ for some $i \in \mathcal{I}$ and $t \geq 0$, then $\Psi_\xi(E) = 1$ if $h^t(h^*) \in E$ and $\Psi_\xi(E) = 0$ otherwise
(ii) if $\xi = I_{i,t}^1\left(\overline{h_i^t}(h^*)\right)$ for some $i \in \mathcal{I}$ and $t \geq 0$, then $\Psi_\xi(E) = 1$ if $\overline{h^t}(h^*) \in E$ and $\Psi_\xi(E) = 0$ otherwise.

Remark 1: Defining an appropriate filtration on the measurable space $(H^\infty, \mathcal{H}^\infty)$, where $\mathcal{H}^\infty$ denotes the Borel sets of H^∞, one can show that conditions (i) and (ii) are equivalent to requiring that the system of beliefs Ψ satisfies Bayes' rule whenever possible.

Firm i's continuation payoff after partial history $x^t \in H^t$ is

$$\pi_{i,t}\left(f_i, \gamma_{-i}\right|\alpha_i, P)\left(x^t\right) \equiv \sum_{k=t}^{\infty} \beta^{k+1-t} \cdot a_{i,t}\left[h^*\left(x^t, f_i, \gamma_{-i}\right)\right].$$

Firm i's expected continuation payoff on $I_{i,t}\left(x_i^t\right)$ is $E_{\Psi_{I_{i,t}\left(x_i^t\right)}}\left[\pi_{i,t}\left(f_i, \gamma_{-i}\right|\alpha_i, P)\right]$.

Since each young worker born at t contacted by firm i observes only firm i's wage offer, the probability of being promoted the following period depends on the number of young workers who accept employment in firm i at date t. Hence, his payoff after partial history $\overline{x^t} \in \overline{H^t}$ is:

$$U_{i,t}\left(\widehat{\sigma}_{i,t}, \gamma\right)\left(\overline{x^t}\right) \equiv \begin{cases} w_{i,t} + \beta \cdot \left(\frac{s_{i,t+1}(h')}{\widehat{l}_{i,t}(h')} \cdot \left[u_{i,t+1}\left(\sigma_{i,t+1}^2, h'\right) - \overline{w}_2\right] + \overline{w}_2\right) & \text{if } \widehat{\sigma}_{i,t}^1\left(w_{i,t}\left(h'\right)\right) = A \\ \overline{w}_1 + \beta \cdot \overline{w}_2 & \text{otherwise} \end{cases}$$

where h' stands for $h^*\left(\overline{x^t}, \widehat{\sigma}_{i,t}, \gamma\right)$. His expected payoff at $I_{i,t}^1(w)$ is $E_{\Psi_{I_{i,t}(w)}}\left[U_{i,t}\left(\widehat{\sigma}_{i,t}, \gamma\right)\right]$.

Finally, I define the equilibrium concept for the game of imperfect but complete information $\Gamma(P)$.

Definition 2.2 *A Perfect Bayesian Equilibrium (PBE) of $\Gamma(P)$ is a profile of strategies $\gamma = \{f_i, \sigma_i\}_{i \in \mathcal{I}}$ and a system of beliefs Ψ such that for every $i \in \mathcal{I}$, $h \in H^\infty$, $t \geq 0$ and realization of $\widehat{v}_t^i(h)$,*

1. $\widehat{u}_{i,t}\left(\sigma_{i,t}^2\right)\left(\overline{v}_{i,t}(h)\right) \geq \widehat{u}_{i,t}\left(\widetilde{\sigma}_{i,t}^2\right)\left(\overline{v}_{i,t}(h)\right)$ for all $\widetilde{\sigma}_{i,t} \in \mathbb{W}_t$.

2. $E_{\Psi_{I_{i,t}^1\left(w_{i,t}(h)\right)}}\left[U_{i,t}\left(\sigma_{i,t}, \gamma\right)\right] \geq E_{\Psi_{I_{i,t}^1\left(w_{i,t}(h)\right)}}\left[U_{i,t}\left(\widetilde{\sigma}_{i,t}, \gamma\right)\right]$ for all $\widetilde{\sigma}_{i,t} \in \mathbb{W}_t$

3. $E_{\Psi_{I_{i,t}\left(h_i^t(h)\right)}}\left[\pi_{i,t}\left(f_i, \gamma_{-i}\left|\alpha_i, P\right)\right] \geq E_{\Psi_{I_{i,t}\left(h_i^t(h)\right)}}\left[\pi_{i,t}\left(\widetilde{f}_i, \gamma_{-i}\right|\alpha_i, P\right)\right]$ for all $\widetilde{f}_i \in \mathbb{F}_i$.

4. Ψ is consistent with γ.

3. Equilibrium Characterization

In this section I characterize the equilibrium wages, vacancies and financial capital. I begin providing a version of the one-stage deviation property for $\Gamma(P)$. In section 3.1, I show that skilled workers are paid their best outside offer. In particular, when no worker changes firms after the training process has ended, competition drives the skilled workers' wage (weakly) above $\frac{c}{\theta}$, the wage that makes a firm indifferent between promoting internally and hiring a worker trained by another firm. In section 3.2, I provide a mild condition on beliefs under which young workers obtain their reservation lifetime utility and explain why this condition implies that firms that promote a larger fraction of their workers pay, *ceteris paribus*, lower entry-wages. In section 3.3, I solve for the vacancies opened by each firm and the associated sequences of entry-wages and growth rates in an equilibrium in which workers get their reservation lifetime utility and firms offer the same wage to externally trained workers.

The equilibrium path of the game $\Gamma(P)$ satisfies a strong version of the one-stage deviation property: if a firm has a short run gain at τ, then it cannot promote $s_{i,\tau+1}(h^*)$ workers the following period. Otherwise the firm could continue with the same hiring policy it chooses on h^* and reinvest the additional assets in bonds forever, increasing its payoff.

Proposition 3.1 *Suppose (γ, Ψ) is a PBE, $i \in \mathcal{I}$ and $\tau \geq 0$. Then $R_{i,\tau}(h) \leq R_{i,\tau}(h^*)$ for any $h \in H^\infty$ such that $h^\tau(h) = h^\tau(h^*)$, $d_{j,\tau}(h) = d_{j,\tau}(h^*)$ for every $j \neq i$, $\lambda \cdot \widehat{l}_{i,\tau}(h) \geq s_{i,\tau+1}(h^*)$ and $h = h^*\left(\overline{h^\tau}(h), \gamma\right)$.*

3.1 The equilibrium wages of skilled workers.

Suppose (γ, Ψ) is a PBE of $\Gamma(P)$ and h^* is the equilibrium path. Proposition 3.2 (i) argues that the wage offer made by a firm where some workers accept an internal promotion must be at least what those workers could make at home. In Proposition 3.2 (ii) I show that if every worker trained by firm i accepts an internal promotion, then firm i's skilled workers are paid their best outside offer. That is, the wage of skilled workers matches either what they can make at home or the offer of some firm j, that is $v^e_{j,t}(h^*) = v_{i,t}(h^*) > \overline{w}_2$; otherwise firm i could offer a slightly lower wage to the workers it trains and increase its earnings.

Proposition 3.2 *If $\widehat{s}_{i,t}(h^*) > 0$, then*
i. $v_{i,t}(h^*) \geq \overline{w}_2$.
ii. If $\widehat{s}_{i,t}(h^*) = s_{i,t}(h^*)$ and $v_{i,t}(h^*) > \overline{w}_2$, then $v_{i,t}(h^*) = v^e_{j,t}(h^*)$ for some $j \in \mathcal{D}_i$.

Trivially, Proposition 3.2 (i) implies that when $\overline{w}_2 \geq \frac{c}{\theta}$, any firm that promotes workers internally offers at least $\frac{c}{\theta}$. Proposition 3.3 shows that if competition for skilled workers drives the wage of internally promoted workers above $\overline{w}_2$ and firms retain every worker they train, the same conclusion holds when $\frac{c}{\theta} > \overline{w}_2$. To see why, notice that since the marginal rate of technical substitution between internally and externally trained workers is $\frac{1}{1+\theta}$, firms are indifferent between the two factors if the wage is $\frac{c}{\theta}$. If the lowest wage offered to internally promoted workers in the industry were smaller than $\frac{c}{\theta}$, some firm i would pay them a wage so close to that wage that every firm $j \in \mathcal{D}_i$ would strictly prefer to hire $(1+\theta)$ workers trained by firm i at a wage slightly above $v_{i,t}(h^*)$ rather than promoting a worker internally. Moreover, by Proposition 3.2 (ii), some firm $j \in \mathcal{D}_i$ would announce a wage $v^e_{j,t}(h^*) = v_{i,t}(h^*)$. If workers trained by firm i did not have any outside offer that matches firm i's offer, firm i could profitably deviate by lowering its wage offer to internally promoted workers. If, instead, some workers trained by firm i rejected firm j's offer, firm j could profitable deviate by raising its wage offer slightly, covering every external vacancy it offers to workers trained by firm i and reducing its internal promotions.

Proposition 3.3 *Suppose $\frac{c}{\theta} > \overline{w}_2$. If $\widehat{s}_{i,t}(h^*) = s_{i,t}(h^*) > 0$ and $v_{i,t}(h^*) > \overline{w}_2$ for every $i \in \mathcal{I}$, then $v_{i,t}(h^*) \geq \frac{c}{\theta}$ for every $i \in \mathcal{I}$.*

Since old workers can obtain $\overline{w}_2$ at home and firms are indifferent between internal promotions and externally trained workers at wage $\frac{c}{\theta}$, intuition suggests that competition for trained workers would drive their best outside offer to $v^* \equiv \max\left\{\overline{w}_2, \frac{c}{\theta}\right\}$. This observation and Proposition 3.3 raises the interest for a particular class of PBE.

Definition 3.1 *A Perfect Bayesian Equilibrium with Symmetric Outside Offers (PBESO) is a PBE (γ, Ψ) such that $v^e_{i,t}(h^*) = v^e_{j,t}(h^*)$ for every $i, j \in \mathcal{I}$. A *PBESO is a PBESO with $v^e_{i,t}(h^*) = v^*$ and $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}\left[U_{i,t}(\sigma_{i,t}, \gamma)\right] = \overline{w}_1 + \beta \cdot \overline{w}_2$ for every $i \in \mathcal{I}$.*

3.2 Entry wages and Prospects for Advancement

In this section I analyze how the entry-wage relates to the firm's promotion policy along the equilibrium path of a PBE. Under a mild condition on beliefs, In Proposition 3.4 I show that young workers get their reservation lifetime utility. In Proposition 3.5 I argue that when the skilled workers' wage is at least v^* and young workers obtain their reservation lifetime utility, assumption AW suffices to rule out the possibility of a shortage in the internal pool of high ability workers . Later, I show that firms that promote a larger fraction of their workers can hire, *ceteris paribus*, workers at a lower entry wage , i.e. they display better prospects for advancement.

Since firms have no incentive to pay young workers more than what is necessary to induce them to accept employment, intuition suggests that the equilibrium entry-wage should be the lowest wage that keeps them indifferent between his two options. A sufficient condition for this result is that the set of wage offers that are strictly preferred to working at home is an open set, a condition that is always satisfied when $\overline{w}_2 \geq \frac{c}{\theta}$ and $v_{i,t}(h) = v^e_{j,t}(h) = \overline{w}_2$ for every $j \in \mathcal{S}_i$ and $h \in H^\infty$ because in that case $E_{\Psi_{I^1_{i,t}(w)}}[U_{i,t}(\sigma_{i,t}, \gamma)] = w + \beta \cdot \overline{w}_2$.

Proposition 3.4 *Suppose* $\left\{w \in \Re_+ : E_{\Psi_{I^1_{i,t}(w)}}[U_{i,t}(\sigma_{i,t}, \gamma)] > \overline{w}_1 + \beta \cdot \overline{w}_2\right\}$ *is an open set. If* $\widehat{l}_{i,t}(h^*) > 0$*, then* $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}[U_{i,t}(\sigma_{i,t}, \gamma)] = \overline{w}_1 + \beta \cdot \overline{w}_2$.

Since in the early stages of the evolution of an industry firms are financially constrained, those that pay lower wages produce more and obtain more revenue to finance expansion. In order to explain the outcome of industry evolution it is important, therefore, to identify what enables one firm to hire workers at a lower wage than another. Insofar worker's abilities are, at least to some degree, firm specific and developed by on-the-job training, one would expect that a young worker's entry-wage depends not only on his opportunity cost and future wages, but also on other factors such as his beliefs about the opportunities for promotion within the firm. For the moment, I will be rather vague and call all those relevant factors "the prospects for advancement" displayed by the firm.

Definition 3.2 *A worker believes that firm i displays better prospect for advancement than firm j if he is willing to work in firm i at a lower wage than in firm j.*

Consider a young worker who receives a wage offer w, believes the probability of being promoted is z and his wage upon promotion will be v^*. Then his lifetime expected utility is $w + \beta \cdot [z \cdot (v^* - \overline{w}_2) + \overline{w}_2]$ if he joins the firm or $\overline{w}_1 + \beta \cdot \overline{w}_2$ if he works at home. Let $w(z, v^*)$ be his *reservation entry-wage*, the wage which makes him indifferent between accepting a job at date t or not. It is the unique w which solves

$$w + \beta \cdot [z \cdot (v^* - \overline{w}_2) + \overline{w}_2] = \overline{w}_1 + \beta \cdot \overline{w}_2 \quad \Leftrightarrow \quad w(z, v^*) = \overline{w}_1 - \beta \cdot z \cdot (v^* - \overline{w}_2)$$

What makes a worker believe one firm displays better prospects for advancement than another? Lemma 3.1 shows that, *ceteris paribus*, one firm displays better prospects for advancement than another at date t if and only if the worker expects the former will promote a larger fraction of its employees than the latter at $t + 1$.

Lemma 3.1 *Suppose* $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}[U_{i,t}(\sigma_{i,t}, \gamma)] = E_{\Psi_{I^1_{j,t}(w_{j,t}(h^*))}}[U_{j,t}(\sigma_{j,t}, \gamma)] = \overline{w}_1 + \beta \cdot \overline{w}_2$. *If* $v_{i,t+1}(h^*) = v_{j,t+1}(h^*) > \overline{w}_2$ *and* $v_{i,t+1}(h^*) = \max_{k \in \mathcal{D}_i \cup \mathcal{D}_j} \left\{v^e_{k,t+1}(h^*)\right\}$*, firm i displays better prospects for advancement at date t on path h^* than firm j if and only if* $\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)} > \frac{s_{j,t+1}(h^*)}{\widehat{l}_{j,t}(h^*)}$.

There are at least two factors that could limit the growth of firms along the process of industry evolution. First, firms might not achieve their optimal size immediately because they do not have enough financial capital to finance expansion. This is represented by the financial capital constraints (2) and (7). Second, firms may face a shortage in their internal labor market. That is, even if financial capital were available to promote more workers, the internal pool might not contain as many high ability candidates as workers the firm would like to hire. This is constraint (6). In principle, any of these constraints may be binding during the process of industry evolution. In Proposition 3.5, however, I show that assumption AW rules out this second possibility along the equilibrium path. Indeed, since a young worker hired by firm i at date t believes his probability of being trained the following period is $\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)}$, the internal labor market constraint does not bind on h^* provided young workers are paid their reservation entry wage and $v_{i,t+1}(h^*) \geq v^*$ upon promotion. Otherwise, assumption AW implies his reservation entry-wage is zero which would contradict the existence of a best response for firm i.

Proposition 3.5 *Suppose AW &* $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}\left[U_{i,t}(\sigma_{i,t},\gamma)\right] = \overline{w}_1 + \beta\overline{w}_2$. *If* $v_{i,t+1}(h^*) \geq v^*$ *then* $\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)} < \lambda$.

3.3 The evolution of entry-wages in a *PBESO.

In this section I characterize the vacancies and entry-wages along the equilibrium path of a *PBESO where skilled workers are paid v^*. I begin with a result I use in Proposition 3.6 (iv) to characterize the relationship between entry-wages and the growth rate of financial capital.

Lemma 3.2 *For any* $g \geq 0$, *the equation* $w = \overline{w}_1 - \beta \cdot \frac{1-\alpha_i}{\alpha_i} \cdot \frac{w}{v^*+c} \cdot g \cdot (v^* - \overline{w}_2)$ *has a unique solution* $\omega : \Re_+ \times (0,1) \mapsto [0, \overline{w}_1]$ *given by* $\omega(g,\alpha_i) = \frac{\alpha_i \cdot (v^*+c) \cdot \overline{w}_1}{\alpha_i \cdot (v^*+c) + (1-\alpha_i) \cdot \beta \cdot (v^* - \overline{w}_2) \cdot g}$. *If* $v^* > \overline{w}_2$, *the function* ω *is continuous and strictly decreasing in* g. *If AW holds,* $\frac{1-\alpha_i}{\alpha_i} \cdot \frac{\omega(g,\alpha_i)}{v^*+c} \cdot g \leq \lambda$ *for any* $g \geq 0$.

In Proposition 3.6 (i) I prove that if wage offers leave young workers indifferent between accepting or rejecting employment, the entry-wage, $w_{i,t}(h^*)$, is $w\left(\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)}, v^*\right)$. In Proposition 3.6 (ii) I show that when $v^* = \frac{c}{\theta}$, old workers do not change firms after the training process has ended. Otherwise, since training is costly and firms are indifferent between promoting internally and hiring externally at wage $\frac{c}{\theta}$, a firm that loses some trained workers would have a profitable deviation. Indeed, that firm could open exactly as many vacancies for externally trained workers as it needs to produce $q_{i,t}(h^*)$ and raise slightly its wage offer to every high ability worker so that those vacancies would be covered and no internally promoted worker would leave after the training process has ended. Since the wage increase could be arbitrarily small, the increase in the wage bill would be more than compensated by the reduction in training costs and the freed financial capital could be invested in bonds. Proposition 3.6 (iii) applies Proposition 3.1 to argue that if (6) does not bind at t and $t+1$ on h^*, the firm hires workers, at wages $w_{i,t}(h^*)$ and v^*, to maximize next period sales revenue subject to a financial constraint. To understand the intuition, notice that if the firm were to deviate from this criterion, it would have fewer assets the following period. Since the firm wants to maximize the discounted sum of its assets and the production function is concave, this deviation could only make sense if it relaxes constraint (6), a possibility that is ruled out by hypothesis. This result together with Proposition 3.5 and assumption AW allows me to write the promotion rate and the young workers' entry-wage as a function of the financial capital growth rate. Indeed, since $\left[\widehat{l}_{i,t}(h^*), s_{i,t}(h^*)\right]$ solves

$$\max_{l,s\geq 0} p_t \cdot l^{\alpha} \cdot s^{1-\alpha} \qquad s.t. \quad w_{i,t}(h^*) \cdot l + (v^*+c) \cdot s = m_{i,t}(h^*) \tag{10}$$

then $\widehat{l}_{i,t}(h^*) = \frac{\alpha_i}{w_{i,t}(h^*)} \cdot m_{i,t}(h^*)$, $s_{i,t}(h^*) = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,t}(h^*)$ and firm i's promotion rate is $\frac{1-\alpha_i}{\alpha_i} \cdot \frac{w_{i,t}(h^*)}{v^*+c} \cdot \frac{m_{i,t+1}(h^*)}{m_{i,t}(h^*)}$.

Proposition 3.6 (iii) establishes that $w_{i,t}(h^*)$ solves

$$w = \overline{w}_1 - \beta \cdot \frac{1-\alpha_i}{\alpha_i} \cdot \frac{w}{v^*+c} \cdot \frac{m_{i,t+1}(h^*)}{m_{i,t}(h^*)} \cdot (v^* - \overline{w}_2) \tag{11}$$

Proposition 3.6 *Suppose* (γ, Ψ) *is a PBE with* $v_{i,t}(h^*) = v^e_{i,t}(h^*) = v^*$ *and* $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}[U_{i,t}(\sigma_{i,t}, \gamma)] = \overline{w}_1 + \beta \cdot \overline{w}_2$ *for every* $i \in \mathcal{I}$ *, then*

i. $w_{i,t}(h^*) = w\left(\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)}, v^*\right)$.

ii. If $v^* = \frac{c}{\theta}$, *then* $\widehat{e}_{i,t}(h^*) = 0$.

iii. If $\frac{s_{i,\tau}(h^*)}{\widehat{l}_{i,\tau-1}(h^*)} < \lambda$ *for* $\tau \in \{t, t+1\}$, *then* $l_{i,t}(h^*) = \frac{\alpha_i}{w_{i,t}(h^*)} \cdot m_{i,t}(h^*)$ *and* $s_{i,t}(h^*) = \frac{(1-\alpha_i)}{v^*+c} \cdot m_{i,t}(h^*)$.

iv. If $\frac{s_{i,\tau}(h^*)}{\widehat{l}_{i,\tau-1}(h^*)} < \lambda$ *for* $\tau \in \{t, t+1\}$, *then* $w_{i,t}(h^*) = \omega\left(\frac{m_{i,t+1}(h^*)}{m_{i,t}(h^*)}, \alpha_i\right)$.

Proposition 3.6 (iv) makes it clear that to obtain the sequence of entry-wages one needs to pin down the sequence of financial capital growth rates and this is the issue I focus on for the reminder of this section. There are two levels of the reservation entry-wage that are key: the reservation entry-wage associated with an stationary level of financial capital, $\omega(1, \alpha_i)$, and the reservation entry-wage associated with a growth rate of r, $\omega(r, \alpha_i)$. For each of these wages, one can define the output price so that the firm's rate of return is r, i.e. the output price which is equal to the firm's marginal cost.

Definition 3.3 *Let* $p_s(\alpha_i) \equiv \frac{r}{q(\omega(1,\alpha_i),\alpha_i)}$ *and* $p_r(\alpha_i) \equiv \frac{r}{q(\omega(r,\alpha_i),\alpha_i)}$.

Remark 2: Since $\omega(1, \alpha_i) > \omega(r, \alpha_i)$ when $v^* > \overline{w}_2$, then $p_s(\alpha_i) > p_r(\alpha_i)$ if $v^* > \overline{w}_2$ and $p_s(\alpha_i) = p_r(\alpha_i) = p^*(\alpha_i)$ if $v^* = \overline{w}_2$.

Suppose at $t+1$ firm i fully reinvests its earnings from sales, that is $\frac{m_{i,t+1}(h^*)}{m_{i,t}(h^*)} = p_t \cdot q(w_{i,t}(h^*), \alpha_i)$, and young workers get their reservation utility. Then, the entry-wage, $w_{i,t}(h^*)$, solves the following equation in w:

$$w = \overline{w}_1 - \beta \cdot \frac{1-\alpha_i}{\alpha_i} \cdot \frac{w}{v^*+c} \cdot p_t \cdot q(w, \alpha_i) \cdot (v^* - \overline{w}_2) \tag{12}$$

Lemma 3.3 *There exists a unique* $\omega^H : \Re_+ \times (0,1) \to [0, \overline{w}_1]$ *that solves (12). If* $v^* > \overline{w}_2$, *the function* $\omega^H(p, \alpha_i)$ *is continuous and strictly decreasing in* p.

In the presence of financial constraints and no fixed costs, it seems reasonable to conjecture that the early stages of the process of industry evolution are characterized by a high output price and positive profits. This induces some firms to fully reinvest their earnings, driving down the output price until profits vanish. It is also natural to think the industry eventually converges to a steady state where financial capital stays constant and every firm makes zero profits. If this conjecture is correct and firms are *ex-ante* identical, Proposition 3.6 implies that the young workers' wage and the output price converge to $\omega(1, \alpha_i)$ and $p_s(\alpha_i)$, respectively. The rate of return of firms that fully reinvest earnings along the transition to the steady state is bounded above by $p_t \cdot q\left(\omega^H(p_t, \alpha_i), \alpha_i\right)$. For profits to be positive along the transition, therefore, it is necessary that $p_t \cdot q\left(\omega^H(p_t, \alpha_i), \alpha_i\right) > r$ or equivalently $p_t > p_r(\alpha_i)$. Thus, it seems natural to restrict the search to sequences of prices in the set

$$\Sigma = \{P \in \Re^\infty : \exists\, T_s \text{ such that } \forall i \in \mathcal{I}, p_t > p_r(\alpha_i) \text{ if } t < T_s \text{ and } p_t = p_s(\alpha_i) \text{ if } t \geq T_s\}$$

Let $\omega^L(p,\alpha)$ be the unique w solving $q(w,\alpha) = \frac{r}{p}$. In Proposition 3.7 I show that for any $P \in \Sigma$ the equilibrium path of a *PBESO where firm i's skilled workers are paid v^* displays the following property: if τ is the first time firm i's entry-wage, $w_{i,t}(h^*)$, differs from $\omega^H(p_t,\alpha_i)$, then it must be strictly greater than $\omega^H(p_\tau,\alpha_i)$ at that date; furthermore, firm i's profits are zero at any date following τ, i.e. $w_{i,t}(h^*) = \omega^L(p_t,\alpha_i)$ for all $t \geq \tau$. Since the entry-wage is different from $\omega^H(p_\tau,\alpha_i)$ at date τ, it must be that firm i does not fully reinvest its earnings as financial capital at $\tau+1$. Because entry-wages are decreasing in financial capital, the date τ entry-wage must be higher than $\omega^H(p_\tau,\alpha_i)$. Since firms maximize revenue subject to a financial constraint, if a firm does not allocate all its assets to financial capital at some date τ, then it must be because it makes zero profits at date τ. The remaining result follows because once maximum profits are zero at some date, they must be zero forever. The intuition behind this is as follows. Since the output price is strictly above p_r, the firm makes zero profits only if young workers born at t believe it will not fully reinvest the sales revenue as financial capital at $t+1$; the latter can only happen if at the $t+1$ entry-wage the firm makes zero profits again.

Proposition 3.7 *Suppose AW holds, $P \in \Sigma$ and (γ,Ψ) is a *PBESO with $v_{i,t}(h^*) = v^* > \overline{w}_2$ for every $t \geq 0$. If there exists τ such that $w_{i,\tau}(h^*) \neq \omega^H(p_\tau,\alpha_i)$ for the first time, then $w_{i,\tau}(h^*) \in \left(\omega^H(p_\tau,\alpha_i), \overline{w}_1\right]$ and $w_{i,t}(h^*) = \omega^L(p_t,\alpha_i)$ for every $t \geq \tau+1$ such that $m_{i,t}(h^*) > 0$.*

3.4 Discussion of assumptions and results

Even though the possibility that a shortage in the firm's internal labor market can be responsible for the slow growth of an industry is realistic, it complicates the analysis enormously. Assumption AW rules that possibility out because it implies that promotion brings about a welfare change that is so large that if workers believed the promotion probability were λ and the wage upon promotion were $\frac{c}{\theta}$, they would work for free when young, i.e. $w\left(\lambda,\frac{c}{\theta}\right) \leq 0$, which would contradict the existence of a best response for the firms.

Propositions 3.6 and 3.7 let one conclude that each *PBESO where $v_{i,t}(h^*) = v^*$ describes the behavior of firms that set up a closed internal labor market with one entry port and an *up or out* promotion system where skilled workers are paid their best outside offer and no trained worker changes firms after the training process has ended.[16] If $\frac{c}{\theta} > \overline{w}_2$, the skilled worker's wage is above what they would obtain out of the industry and so small changes in market conditions (i.e. changes in $\overline{w}_2$) do not affect it.

Moreover, Propositions 3.5 and 3.6 (iii) imply that if AW holds, each firm maximizes next period's sales revenue subject to a financial constraint, firm i's production is $q(w_{i,t}(h^*),\alpha_i)\cdot m_{i,t}(h^*)$, where $q(w_{i,t}(h^*),\alpha_i) \equiv \left(\frac{\alpha_i}{w_{i,t}(h^*)}\right)^{\alpha_i}\cdot\left(\frac{1-\alpha_i}{v^*+c}\right)^{1-\alpha_i}$, and the financial capital rate of return at t is $R^*(p_t,w_{i,t}(h^*),\alpha_i) \equiv p_t\cdot q(w_{i,t}(h^*),\alpha_i)$.

Finally, Propositions 3.6 (iv) and 3.7 let us conclude that if AW holds, either each firm makes zero profits and fully reinvests its capital up to date $T-1$ or it makes zero profits and its financial capital growth rate is $\omega^{-1}\left(\omega^L(p_t,\alpha_i),\alpha_i\right)$ from date 1 on, where ω^{-1} is the inverse of ω

[16] The best outside offer depends on c and θ because of the assumption that some firms are yet to train when others have ended their training process. This contrasts with Bernhardt and Scoones [6] where competitors incur a cost to learn the ability of outside workers but that cost does not affect the best outside offer. This is because, in their model, firms bid for workers after that cost is sunk.

4. OPTIMISTIC AND PESSIMISTIC PBE

In this section I prove that there exists a *PBESO where skilled workers are paid v^* and *ex ante* identical firms follow different growth paths that converge to a steady state where profits are zero and workers receive the same wage regardless of the firm that employs them. First, I define a family of strategies for firms and workers. Proposition 4.1 shows that any of the sequences of entry-wages, and its associated sequences of financial capital growth rates, described in Proposition 3.7 can be sustained as part of the path of play of a strategy profile in that family. Proposition 4.2 shows that there exists a profile of strategies in that family and a system of beliefs that constitutes a *PBESO for a large subset of output price sequences in Σ.

Let $\widehat{v}^o_{i,t}(h) \equiv \max\left\{\max_{j\in\mathcal{D}_i} \widehat{v}^e_{j,t}, \overline{w}_2\right\}$ be the best outside offer received by a worker trained by firm i who observes a realization $\left\{\widehat{v}^e_{j,t}\right\}_{j\in\mathcal{D}_i}$ on path h. Consider the strategy

$$\widetilde{\sigma}^2_{i,t}\left[\overline{v}_{i,t}(h)\right] = \begin{cases} i & \text{if } v_{i,t}(h) \geq \widehat{v}^o_{i,t}(h) \text{ and } O_{i,t}\left(\overline{v}_{i,t}(h)\right) = \varnothing \\ j & \text{if } \widehat{v}^e_{j,t} = \widehat{v}^o_{i,t}(h) \text{ and } \left(j \in O_{i,t}\left(\overline{v}_{i,t}(h)\right) \text{ or } \widehat{v}^o_{i,t} > v_{i,t}(h)\right) \\ o & \text{otherwise} \end{cases}$$

where $O_{i,t}\left(\overline{v}_{i,t}(h)\right) = \left\{j \in \mathcal{D}_i : \widehat{v}^e_{j,t}(h) = v_{i,t}(h), \chi_{j,t}(h) = 1\right\}$ is the set of firms that match firm i's offer and face a binding internal labor market constraint at information set $I_{i,t}\left(h^t_i(h)\right)$.[17] Old workers choose to work where they obtain the highest wage and when indifferent between two or more wage offers they only leave the place where they were trained if the outside offer comes from a firm which faces a binding internal labor market constraint. Clearly, $\widetilde{\sigma}^2_{i,t}$ satisfies condition (i) in the definition of a PBE.

For each $P \in \Sigma$ and $t \geq 0$ define a family, parameterized by $\delta \in [0,1]$ and $T \geq T_s$, of functions $\widetilde{\sigma}^1_t(T,\delta\,|\alpha,P) : \Re_+ \mapsto \Re_+$ and $\widetilde{f}_t(T,\delta\,|\alpha,P) : \Re^3_+ \mapsto \Re^7_+$ such that $\widetilde{f}_t(T,\delta\,|\alpha,P) = \left(\widetilde{l}_t(\alpha), \widetilde{e}_t(\alpha), \widetilde{w}_t(T,\delta\,|\alpha,P), v^*, v^*, \widetilde{m}_t(T,\delta), \widetilde{b}_t\right)$ where $\widetilde{b}_t(l,m,a) = a - \widetilde{m}_t(T,\delta)$ and

$$\widetilde{\sigma}^1_t(T,\delta\,|\alpha,P)(w) = \begin{cases} A & \text{if } w \geq \omega^H(p_t,\alpha) \text{ and } t < T-1 \\ A & \text{if } w \geq \omega^H(\delta\cdot p_{T-1},\alpha) \text{ and } t = T-1 \\ A & \text{if } w \geq \omega(1,\alpha) \text{ and } t \geq T \\ R & \text{otherwise} \end{cases}$$

$$\widetilde{m}_t(T,\delta)(l,m,a) = \begin{cases} a & \text{if } t < T \\ \delta\cdot a & \text{if } t = T, \\ \max\{m,a\} & \text{if } t > T \end{cases} \qquad \widetilde{w}_t(T,\delta\,|\alpha,P)(l,m,a) = \begin{cases} \omega^H(p_t,\alpha) & \text{if } t < T-1 \\ \omega^H(\delta\cdot p_{T-1},\alpha) & \text{if } t = T-1, \\ \omega(1,\alpha) & \text{if } t \geq T \end{cases}$$

$$\widetilde{l}_t(\alpha)(l,m,a) = \frac{\alpha\cdot\widetilde{m}_t(T,\delta\,|\alpha,P)}{\widetilde{w}_t(T,\delta\,|\alpha,P)}, \quad \widetilde{e}_t(\alpha)(l,m,a) = \begin{cases} \frac{(1-\alpha)\cdot\widetilde{m}(T,\delta)}{v^*} & \text{if } \frac{(1-\alpha)\cdot\widetilde{m}(T,\delta)}{v^*\cdot(1+\theta)} < \lambda\cdot l \\ \frac{(1-\alpha)\cdot\widetilde{m}(T,\delta)}{v^*} - (1+\theta)\cdot\lambda l & \text{otherwise} \end{cases}$$

Define strategies $\sigma^H_i \in \mathbb{W}_i$ and $f^H_i \in \mathbb{F}_i$ as

$$\begin{aligned} \sigma^H_{i,t} &= \left[\widetilde{\sigma}^1_t(T,\delta\,|\alpha_i,P), \widetilde{\sigma}^2_{i,t+1}\right] \\ f^H_{i,t}\left(x^t_i\right) &= \widetilde{f}_t(T,\delta\,|\alpha_i,P)\left(l_{i,t-1}(h), m_{i,t-1}(h), a_{i,t}(h)\right) \quad \text{where } h^t_i(h) = x^t_i \end{aligned}$$

[17] If two firms meet the criteria in the second line of the definition of the strategy, then the worker chooses the one with the lowest subindex. This choice of a tie breaking rule is, of course, without loss of generality.

The cut-off wage of $\sigma_{i,t}^{H}$ is the reservation entry-wage of a worker who believes his promotion probability is that of a firm that reinvests all its assets as financial capital if $t \leq T-1$, it reinvests a fraction δ of them if $t = T$ and keeps its financial capital constant if $t > T$. Firm i offers the cutoff value of $\sigma_{i,t}^{H}$ to the young workers it contacts and spends its financial capital as if it were maximizing short run constrained profits. For fixed $T \geq T_s$ and $\delta \in [0,1]$, $\mathcal{I}^H \subset \mathcal{I}$ denotes the subset of firms where workers and firms play $\left(f_i^H, \sigma_i^H\right)$.

The strategy $\left(f_i^H, \sigma_i^H\right)$ describes the behavior of firm i and the workers when the latter are optimistic about the prospects for advancement displayed by firm i. When $\frac{c}{\theta} > \overline{w}_2$, there are also strategies in which young workers are pessimistic about the prospects for advancement displayed by firm i. For each $w_0^L \in (\underline{w}, \overline{w})$, where $\underline{w} \equiv \omega^H(p_0, \alpha)$ and $\overline{w} \equiv Min\left\{\omega^L(p_0,\alpha), \overline{w}_1\right\}$, define

$$\widetilde{\sigma}_t^1\left(w_0^L \,|\alpha, P\right)(w) = \begin{cases} A & \text{if } w \geq w_0^L \text{ and } t=0 \text{ or } w \geq \omega^L(p_t,\alpha) \text{ and } t \geq 1 \\ R & \text{otherwise} \end{cases}$$

and $\widetilde{f}_t\left(w_0^L \,|\alpha, P\right) : \Re_+^3 \mapsto \Re_+^7$ as $\widetilde{f}_t\left(w_0^L \,|\alpha, P\right) = \left(\widetilde{l}_t(\alpha), \widetilde{e}_t(\alpha), \widetilde{w}_t\left(w_0^L \,|\alpha, P\right), v^*, v^*, \widetilde{m}_t\left(w_0^L \,|\alpha, P\right), \widetilde{b}_t\right)$ where $G(w,\alpha)$ is the inverse of $\omega(g,\alpha)$ and

$$\widetilde{m}_t\left(w_0^L \,|\alpha, P\right)(l,m,a) = \begin{cases} a & \text{if } t=0 \\ \min\left\{G\left(w_0^L,\alpha\right)\cdot a, a\right\} & \text{if } t=1, \\ \min\left\{G\left[\omega^L(p_t,\alpha),\alpha\right] m, a\right\} & \text{if } t \geq 2 \end{cases} \quad \widetilde{w}_t\left(w_0^L \,|\alpha, P\right) = \begin{cases} w_0^L & \text{if } t=0 \\ \omega^L(p_t,\alpha) & \text{if } t \geq 1 \end{cases}$$

For fixed $w_0^L \in (\underline{w}, \overline{w})$ define strategies $\sigma_i^L \in \mathbb{W}_i$ and $f_i^L \in \mathbb{F}_i$ as

$$\begin{aligned} \sigma_{i,t}^L &= \left[\widetilde{\sigma}_t^1\left(w_0^L \,|\alpha_i, P\right), \widetilde{\sigma}_{i,t+1}^2\right] \\ f_{i,t}^L\left(x_i^t\right) &= \widetilde{f}_t\left(w_0^L \,|\alpha_i, P\right)\left(l_{i,t-1}(h), m_{i,t-1}(h), a_{i,t}(h)\right) \quad \text{where } h_i^t(h) = x_i^t \end{aligned}$$

If firm i faces workers that play strategy σ_i^L, it can just make zero profits from date 1 on hiring workers at the cutoff wages and, therefore, it is (weakly) optimal to reduce its financial capital along time; strategy f_i^L specifies a path of reinvestment that justifies the workers' pessimism about the prospects for advancement offered firm i. For fixed $w_0^L \geq 0$, $\mathcal{I}^L \subset \mathcal{I}$ denotes the subset of firms where workers and firms play $\left(f_i^L, \sigma_i^L\right)$.

Let $\widehat{\gamma}$ be the profile of strategies where a fraction μ^H of the firms belongs to $\mathcal{I}^H$ and a fraction $1-\mu^H$ belongs to $\mathcal{I}^L$. The following proposition shows that on the path of play $\widehat{h}$ induced by $\widehat{\gamma}$, no trained worker change firms and no firm face a binding internal labor market constraint. In addition, firms in $\mathcal{I}^H$ fully reinvest earnings while firms in $\mathcal{I}^L$ reduce their financial capital during the transition towards the steady state.

Proposition 4.1 *Suppose $P \in \Sigma$ and $\widehat{h} \equiv h^*(\widehat{\gamma})$. Then,*

i. $\widehat{l}_{i,t}\left(\widehat{h}\right) = \frac{\alpha_i}{w_{i,t}\left(\widehat{h}\right)} \cdot m_{i,t}\left(\widehat{h}\right)$ *for every $t \geq 0$.*

ii. If AW holds, then $\widehat{e}_{i,t}\left(\widehat{h}\right) = 0$ and $s_{i,t}\left(\widehat{h}\right) = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,t}\left(\widehat{h}\right) < \lambda \cdot \widehat{l}_{i,t-1}\left(\widehat{h}\right)$ for every $t \geq 0$. The same conclusion holds if $\frac{c}{\theta} = \overline{w}_2$, $p_0 \leq \frac{p^*(\alpha_i)}{r} \cdot \frac{\alpha_i}{1-\alpha_i} \cdot \frac{\overline{w}_2 + c}{\overline{w}_1} \cdot \lambda$ and P is non-increasing.

iii. If $i \in \mathcal{I}^H$, then $m_{i,t}\left(\widehat{h}\right) = a_{i,t}\left(\widehat{h}\right)$ for every $t \leq T-1$ and $m_{i,t}\left(\widehat{h}\right) = \delta \cdot a_{i,T}\left(\widehat{h}\right)$ for every $t \geq T$.

If $i \in \mathcal{I}^L$, then $m_{i,t}\left(\widehat{h}\right) = a_0$ and $m_{i,t}\left(\widehat{h}\right) = G\left[\omega^L(p_{t-1},\alpha_i),\alpha_i\right] \cdot m_{i,t-1}\left(\widehat{h}\right)$ for every $t \geq 1$.

To show that no trained worker leaves firm i when the skilled worker wage is v^*, it is key to argue that $\chi_{j,t}\left(\widehat{h}\right)=0$ for every $j\in\mathcal{D}_i$; that is the role of the assumptions in Proposition 4.1 (ii). If $\frac{c}{\theta}>\overline{w}_2$, Assumption AW is sufficient because it guarantees the promotion rate is smaller than λ. If $\frac{c}{\theta}=\overline{w}_2$, the upper bound on p_0 implies $\chi_{j,0}\left(\widehat{h}\right)=0$ even for firms that fully reinvest its earnings as financial capital at date 1; since P is a nonincreasing sequence and the financial capital growth rate increases with prices, then $\chi_{j,t}\left(\widehat{h}\right)=0$ for all $t\geq 1$.

To define a PBE one needs to specify the beliefs of the players on information sets off the equilibrium path. Let $\widehat{\Psi}$ be a system of beliefs consistent with $\widehat{\gamma}$ that at each information set, $I_{i,t}\left(x_i^t\right)$ or $I_{i,t}^1(w)$, off the equilibrium path puts mass one on some partial history $\widehat{x}^t\in\Omega_{i,t}\left(x_i^t\right)$ or $\widehat{\overline{x^t}}\in\Omega_{i,t}^1(w)$, respectively.[18] $\Omega_{i,t}\left(x_i^t\right)\subset I_{i,t}\left(x_i^t\right)$ is the set of paths where (a) firm i's competitors would not face a binding internal labor market constraint even if all its external offers were rejected, and (b) they offer v^* to trained workers, as on $\widehat{h}$, and (c) the supply of externally trained workers i faces exceeds i's demand, what is consistent with perfect competition in the output market. Off the equilibrium path, therefore, firm i believes there are enough trained workers willing to join it and, therefore, (6) limits only the number of internal promotions but not its production level, what allows me to break the intertemporal problem of the firm at each information set in a sequence of one-period problems.[19] For the young worker, I assume that when $w>w_{i,t}\left(\widehat{h}\right)$, $\Omega_{i,t}^1(w)\subset I_{i,t}^1(w)$ is the set of paths where firm i's behavior differs from $d_{i,t}\left(\widehat{h}\right)$ in that it opens more external vacancies, a condition that guarantees the promotion rate and, therefore, the young worker's expected utility increases with wages. If $w<w_{i,t}\left(\widehat{h}\right)$, the young worker's payoff is independent of beliefs because he anticipates every other young worker rejects employment so that nobody will be promoted at $t+1$. Hence, I assume $\Omega_{i,t}^1(w)=I_{i,t}^1(w)$. So, $\widehat{\Psi}$ seems a natural choice for the players' beliefs.

Let $p^{**}(\alpha_i)=\frac{r}{q(\overline{w}_1,\alpha_i)}$ be the marginal cost of a firm expected to shut down the next period, i.e. $\overline{w}_1=\omega(0,\alpha_i)$.

Proposition 4.2 *Suppose $P\in\Sigma$, $\delta\in[0,1]$ and $p_{T-1}\cdot q\left(\omega^H\left(\delta\cdot p_{T-1},\alpha_i\right),\alpha_i\right)>r$ for every $i\in\mathcal{I}^H$.*

*i. If $\mu^H=1$ and AW holds, then $\left(\widehat{\gamma},\widehat{\Psi}\right)$ is a *PBESO. The same conclusion holds if $\mu^H=1$, $\frac{c}{\theta}=\overline{w}_2$, $p_0\leq\frac{p^*(\alpha_i)}{r}\cdot\frac{\alpha_i}{1-\alpha_i}\cdot\frac{\overline{w}_2+c}{\overline{w}_1}\cdot\lambda$ for every $i\in\mathcal{I}$ and P is non-increasing.*

*ii. If $\mu^H<1$, AW holds, $w_0^L\in(\underline{w},\overline{w})$, and $p_t\leq p^{**}(\alpha_i)$ for every $t\geq 1$ and $i\in\mathcal{I}^L$, then $\left(\widehat{\gamma},\widehat{\Psi}\right)$ is a *PBESO.*

The *PBESO in (i) describes an industry where workers are optimistic about the prospects for advancement displayed by every firm. In the PBE described in (ii), workers are optimistic about some firms and pessimistic about others and this may happen even if firms are *ex-ante* identical, that is when $\alpha_i=\alpha$ for every $i\in\mathcal{I}$. Here one needs to impose the additional assumption that $p_t\leq p^{**}(\alpha_i)$ for every $t\geq 1$ and $i\in\mathcal{I}^L$ so that the growth rate of firms in $\mathcal{I}^L$ is nonnegative. The behavior of firms in $\mathcal{I}^H$ or $\mathcal{I}^L$ differs along the transition to the steady state. Even though *ex-ante* identical firms follow different growth paths, the worker's entry-wage converges to the same level at date T because $w_{i,t}(h)=\omega^L(p_s,\alpha_i)=\omega(1,\alpha_i)$ for every $t\geq T$. Hence, firms playing f_i^H and f_i^L stop

[18] A formal definition of these sets can be found in the appendix.
[19] Recall that when $\frac{c}{\theta}\geq\overline{w}_2$, it costs the same to produce one unit of task 2 employing internally promoted workers or hiring externally trained workers.

growing and pay the same wage from date T on. Any difference in their steady state size, therefore, originates during the transition towards the stationary state.

For fixed μ^H, Proposition 4.2 identifies a continuum of PBE indexed by $\delta \in [0, 1]$ and $w_0^L \in (\underline{w}, \overline{w})$. In section 5, I show that for each $w_0^L \in (\underline{w}, \overline{w})$ the value of δ is pinned down by the output market clearing condition.

4.1 Robustness

The assumption that workers live for two periods, together with the fact that the internal labor market constraint never binds, implies firms never raid and train immediately a high ability worker who is not promoted by other firm. This is because he is less productive than a worker promoted internally but must be paid the same wage upon promotion. If the worker lived for three periods or more, however, he could be trained in the third period of his life either by the first period employer or by a firm who raid him in the second period to perform the first task. However, no firm would be willing to do so because it costs the same to train him than to train a worker who has developed high ability and is two-years old but the latter can perform the second task for one period more than the former. Consequently, a high ability worker who has not been trained by another firm can perform only task 1 and must be paid at least $\overline{w}_2$. Since young workers are willing to perform the same task at a lower wage, therefore, no firm would raid high ability workers even if they lived more than two periods. Likewise, an increase in the number of periods a worker lives does not affect his wage upon promotion since this wage is determined by the trade-off between the training cost and the specific ability faced by the competitors of his first-period employer.

5. INDUSTRY EQUILIBRIUM

The concept of equilibrium used in the previous sections does not require the output market to clear at the prices the firms take as given. This is a drawback because the evolution of output prices is related to the evolution of firms' assets through the market clearing condition. To capture this aspect of industry evolution, I define an Industry Equilibrium (IE) as a collection of strategies, beliefs and output prices P such that the strategies of firms and workers are a PBE of $\Gamma(P)$ and the output market clears on the equilibrium path of $\Gamma(P)$.

To simplify the exposition, I assume $\alpha_i \in \{\alpha_H, \alpha_L\}$ for every $i \in \mathcal{I}$. Moreover, since the family of strategies considered in section 4 is rich enough to sustain any outcome described in Propositions 3.6 and 3.7, I restrict the analysis to that type of strategies. I assume each firm belongs either to $\mathcal{I}^L$ or $\mathcal{I}^H$ with the understanding that if firms are heterogeneous, firms with $\alpha_i = \alpha_H$ are in $\mathcal{I}^H$ and those with $\alpha_i = \alpha_L$ are in $\mathcal{I}^L$; if firms are homogenous, instead, I assume a fraction μ^H belongs to $\mathcal{I}^H$ while the rest are in $\mathcal{I}^L$. This equilibrium concept does not restrict *ex-ante* identical firms to play the same strategy. Heterogeneous behavior may arise either because firms have different technologies or due to a coordination problem among the infinite generations of workers.[20]

[20] A plausible story of how *ex-ante identical* firms end up partitioned in these two sets is as follows. At date zero, after firms announce their names, each worker who is contacted by some employer observes the realization of a binary *sunspot* variable that assigns probability μ^H to f_i^H and $1-\mu^H$ to f_i^L for fixed T, δ and w_0^L and updates his common prior about the strategy of that firm. The realization of the sunspot at date zero induces a decision rule for each generation born at $t \geq 0$, a mapping from the set of firms, $\mathcal{I}$, to the set $\left\{\sigma_{i,t}^L, \sigma_{i,t}^H\right\} \in \mathbb{W}_t \times \mathbb{W}_t$.

Definition 5.1 *An Industry Equilibrium (IE) is a* $\left\{P, \mu^H, \gamma, \Psi\right\}$ *such that* $P \in \Re_+^{\infty}$, $\mu^H \in [0,1]$ *and*

1. (γ, Ψ) *is a *PBESO of* $\Gamma(P)$ *in which either* $i \in \mathcal{I}^L$ *or* $i \in \mathcal{I}^H$,

2. $q_t^H(h^*) \cdot \mu^H + q_t^L(h^*) \cdot \left(1 - \mu^H\right) = D(p_t)$, *for every* $t \geq 0$,

where $q_t^L(h^*)$ *and* $q_t^H(h^*)$ *are the aggregate output produced at date* t *on* h^* *by firms in* $\mathcal{I}^L$ *and* $\mathcal{I}^H$, *respectively.*

In a steady state with zero profits and *ex-ante* identical firms, the entry-wage is $\omega(1, \alpha)$ and the output price is $p_s(\alpha)$ every period. Then, the market clears if and only if the aggregate financial capital is $\frac{p_s(\alpha) \cdot D(p_s(\alpha))}{r}$. For $k \in \{L, H\}$, let $a_t^k(h)$ and $m_t^k(h)$ be the aggregate assets and financial capital of firms in $\mathcal{I}^k$ at t on $h \in H^{\infty}$.

6. Industry Equilibrium with Ex Ante Identical Firms

In this section, I consider an industry with *ex-ante* identical firms, i.e. $\alpha_i = \alpha$ for all $i \in \mathcal{I}$.[21] In section 6.1, I analyze the benchmark case where firms are not financially constrained at date zero. Section 6.2, turns to the more interesting scenario where firms are constrained. In section 6.2.1, I state conditions under which an IE where *ex-ante* identical firms follow identical growth paths exists and is unique; furthermore I show it converges to a Walrasian-like state in finite time and I analyze its efficiency properties. In section 6.2.2, I show, by means of an example, that there is also IE where *ex-ante* identical firms follow different growth paths.

6.1 Unconstrained IE

I begin with the case in which $a_0 \geq \frac{p_s \cdot D(p_s)}{r}$ and so firms have enough assets to drive profits to zero from the start. Let P_s be the sequence with $p_t = p_s$ for all $t \geq 0$. Clearly, $P_s \in \Sigma$. In any IE associated to P_s, every firm is in $\mathcal{I}^H$, for $T = 0$ and $\delta = \frac{p_s \cdot D(p_s)}{r \cdot a_0}$, and the industry output level is $D(p_s)$. Hence, *ex-ante* identical firms produce the same, firms are not financially constrained and workers performing the same task receive the same wage regardless of the firm that employs them. If $v^* = \overline{w}_2$, then $p_s = p^*$ and allocative efficiency holds. Otherwise, too little is produced with respect to the efficient allocation (i.e. $p_s > p^*$) but technological efficiency holds because each firm pays wages $\omega(1)$ and v^* and maximizes profits.

To understand why too little is produced when $v^* > \overline{w}_2$, notice that the marginal cost is $p_s = \frac{r}{q(\omega(1))}$ while the marginal cost in the efficient allocation is $p^* = \frac{r}{q^*}$, where $q^* \equiv \left(\frac{\alpha}{\overline{w}_1}\right)^{\alpha} \cdot \left(\frac{1-\alpha}{\overline{w}_2 + c}\right)^{1-\alpha}$ solves

$$Max\ l^{\alpha} \cdot s^{1-\alpha} \qquad s.t.\ \overline{w}_1 \cdot l + (\overline{w}_2 + c) \cdot s \leq 1 \tag{13}$$

The two marginal costs are equal if and only if $v^* = \overline{w}_2$ and so $\omega(1) = \overline{w}_1$. So it suffices to argue that $v^* > \overline{w}_2$ implies $q(\omega(1)) < q^*$. The answer is not obvious because as v^* increases, the wage of young workers decreases. However, since workers discount the future, a marginal increase in v^* leads to a less than proportional reduction in $\omega(1)$. Since $\omega(1) = \overline{w}_1 - \beta \cdot \frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)} \cdot (v^* - \overline{w}_2)$ and $s_{i,t}(h^*) = s_{i,t+1}(h^*)$, it follows that

$$\begin{aligned} 1 &= \omega(1) \cdot \widehat{l}_{i,t}(h^*) + (v^* + c) \cdot s_{i,t}(h^*) \\ &= \overline{w}_1 \cdot \widehat{l}_{i,t}(h^*) + [(1-\beta) \cdot v^* + \beta \cdot \overline{w}_2 + c] \cdot s_{i,t}(h^*) > \overline{w}_1 \cdot \widehat{l}_{i,t}(h^*) + (\overline{w}_2 + c) \cdot s_{i,t}(h^*) \end{aligned}$$

[21] In this section, to simplify notation, I omit the parameter α in the functions q, ω , ω^H, ω^L, p_s, p_r and p^*.

Therefore, $(l_{i,t}(h^*), s_{i,t}(h^*))$ is in the feasible set of problem (13) which implies that $q(\omega(1)) < q^*$.

Remark 3: This argument shows that the lack of allocative efficiency in the steady state does not depend on the assumption that firms are financially constrained; the impossibility of enforcing a long term contract in which an old worker is paid $\overline{w}_2$ induces the lack of allocative efficiency.

Remark 4: Since the industry is in steady state from the start when $a_0 \geq \frac{p_s \cdot D(p_s)}{r}$, one cannot address issues such as how the prospects for advancement displayed by the firm affects its growth path or its long run size.

6.2 Constrained IE: $0 < a_0 < \frac{p_s \cdot D(p_s)}{r}$

This section analyzes the case in which $0 < a_0 < \frac{p_s \cdot D(p_s)}{r}$ and so the initial financial capital falls short of the steady state level. Proposition 6.1 in section 6.2.1 shows that there is a unique IE where every firm displays identical prospects for advancement and output prices are in Σ. In that equilibrium, firms and workers behave according to the optimistic strategies described in section 4. In section 6.2.2, I show by example that if AW holds, there exists IE in which *ex-ante* identical firms follow different growth paths.

6.2.1 Identical Growth Paths

Suppose there is an IE with $\mu^H = 1$. From date T on, the output price is p_s, the entry-wage is $\omega(1)$, total supply is $q[\omega(1)] \cdot m_T^H(h^*) = \frac{r}{p_s} \cdot m_T^H(h^*)$ and $m_T^H(h^*) = \frac{p_s \cdot D(p_s)}{r}$ by market clearing. The unknowns, then, are the output prices along the transition towards the steady state, that is, the sequence $\{p_t\}_{t=0}^{T-1}$. Since $q_{T-1}^H(h^*) = q[w_{i,T-1}(h^*)]$, $w_{i,T-1}(h^*) = \omega\left(\frac{m_T^H(h^*)}{m_{T-1}^H(h^*)}\right)$, then p_{T-1} is the value of p that solves:

$$q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot m_{T-1}^H(h^*)}\right)\right] \cdot m_{T-1}^H(h^*) = D(p) \tag{14}$$

Proposition 4.1 (iii) implies $\delta = \frac{m_T^H(h^*)}{p_{T-1} \cdot q[w_{i,T-1}(h^*)] \cdot m_{T-1}^H(h^*)}$ and by market clearing one obtains $\delta = \frac{p_s \cdot D(p_s)}{r \cdot p_{T-1} \cdot D(p_{T-1})}$.

If $T = 1$, then (14) completely describes the output prices along the transition to the steady state. If $T > 1$, however, market clearing requires p_t to be the value of p that solves

$$q\left[\omega^H(p)\right] \cdot m_t^H(h^*) = D(p) \qquad \text{for all } 0 \leq t \leq T-2 \tag{15}$$

Lemma 6.1 *Suppose AD holds. If $m \leq \frac{p_s \cdot D(p_s)}{r}$, the equation $q\left[\omega^H(p)\right] \cdot m = D(p)$ has a unique solution $\mathbb{P} : \left[0, \frac{p_s \cdot D(p_s)}{r}\right] \to \left(\frac{p_s}{r}, \infty\right)$ and $\mathbb{P}(m) > p_r$ if and only if $m < \frac{p_r \cdot D(p_r)}{r}$.*

If $T > 1$, Lemma 6.1 implies that $p_0 = \mathbb{P}(a_0)$. One concludes that for any $T \geq 1$, the workers' date zero entry-wage is unique. Indeed, $w_{i,0}(h^*) = \omega\left(\frac{p_s \cdot D(p_s)}{r \cdot a_0}\right)$ if $T = 1$ and $w_{i,0}(h^*) = \omega^H(\mathbb{P}(a_0))$ if $T > 1$. In addition, conditions (14) and (15) show that prices and entry-wages depend only on the aggregate financial capital at any date $t \leq T-1$. Since $m_{i,t}(h^*) = p_{t-1} \cdot q(w_{i,t-1}(h^*)) \cdot m_{i,t-1}(h^*)$ for all $1 \leq t \leq T-1$ and $i \in \mathcal{I}$, the uniqueness of p_0 implies there is at most one equilibrium sequence of prices in Σ when every firm belongs to $\mathcal{I}^H$.

In order to prove the existence of an IE, I construct a sequence $P \in \Sigma$ by iterating the map $\mathbb{P}$ until the first date that full reinvestment of revenues would make the aggregate financial capital larger than $\frac{p_s \cdot D(p_s)}{r}$. That date is the candidate for date $T-1$; to complete the sequence $\{p_t\}_{t=0}^{T-1}$, I choose p_{T-1} to be the solution of (14) given the value of $m_{T-1}^H(h^*)$.[22] Let $\underline{a} = \frac{p^*}{r} \cdot D\left(\frac{p^*}{r}\frac{\alpha}{1-\alpha} \cdot \frac{\overline{w}_2+c}{\overline{w}_1} \cdot \lambda\right)$ be the initial capital such that constraint (6) binds at the date zero output clearing price when $\frac{c}{\theta} = \overline{w}_2$.

Proposition 6.1 *Assume AD. If AW holds or $\frac{c}{\theta} = \overline{w}_2$ and $a_0 \geq \underline{a}$, a unique IE with $P \in \Sigma$ and $\mu^H = 1$ exists.*

In this IE, technological efficiency holds since firms pay the same to workers performing identical tasks. The credit constraint causes a failure of allocative efficiency in the transition to the steady state. If $\frac{c}{\theta} = \overline{w}_2$, however, the efficient output level is achieved in steady state because $D(p_s) = D(p^*)$. If $\frac{c}{\theta} > \overline{w}_2$, instead, allocative efficiency fails even in the steady state because $D(p_s) < D(p^*)$. Since allocative efficiency would also have failed in steady state if firms were unconstrained at date zero but would hold if $\frac{c}{\theta} \leq \overline{w}_2$, one concludes it is the impossibility of enforcing a wage $\overline{w}_2$ for the skilled workers, rather than the credit constraint, what causes industry output to fall short of the efficient level in the long run.

The example in the next section, in turn, shows that a credit constraint is necessary to explain why *ex-ante* identical firms might follow different growth paths and have different market shares in the long run.

6.2.2 Different Growth Paths

Now suppose $\mu^H \in (0,1)$ and AW holds. Proposition 6.2 explains how the retained earnings dynamic selects among firms that promote workers internally since it shows that in any IE with $P \in \Sigma$, those firms that display the worst prospects for advancement at date zero continue to show the worst prospects up to date T.

Proposition 6.2 *Suppose $\mu^H \in (0,1)$ and AW holds. If $w_{i,0}(h^*) \in (\underline{w}, \overline{w})$ and $P \in \Sigma$, then $i \in \mathcal{I}^L$ and $w_{i,t}(h^*) > w_{j,t}(h^*)$ for every t such $0 \leq t \leq T-1$ and $j \in \mathcal{I}^H$.*

It follows from Proposition 6.2 that firms that display better prospects for advancement at date zero have a higher growth rate along the transition towards the steady state and a higher steady state market share than firms that, *ceteris paribus*, initially show worse prospects for advancement. One concludes that among *ex-ante* identical firms, the retained earning dynamic favors those firms that display better prospects for advancement at date zero.

The rest of this section characterizes an IE in which *ex-ante* identical firms follow different growth paths and $T > 1$; later I use this characterization to construct an example. At date zero, the output price must satisfy

$$q\left[\omega^H(p_0)\right] \cdot a_0^H + q\left(w_0^L\right) \cdot a_0^L = D(p_0) \qquad \underline{w} < w_0^L \leq \overline{w} \tag{16}$$

where the left hand side is the date zero short run industry supply function. For any initial level of aggregate financial capital, $a_0 = a_0^H + a_0^L$, the assumption that $w_0^L > \underline{w}$ implies the industry supply shifts to the left when compared to the case in which firms display equal prospects for advancement. Therefore, there is an excess of demand at $\mathbb{P}(a_0)$, the price which solves (16) when $\mu^H = 1$. Likewise, there is an excess of supply at the price $\mathbb{P}\left(a_0^H\right)$. Since the date zero supply function is strictly increasing in prices, demand is strictly decreasing and both functions are continuous, there exists a unique $\widetilde{\mathbb{P}}\left(a_0^H, a_0^L, w_0^L\right)$ that solves (16). In addition, $\mathbb{P}\left(a_0^H + a_0^L\right) < \widetilde{\mathbb{P}}\left(a_0^H, a_0^L, w_0^L\right) < \mathbb{P}\left(a_0^H\right)$. It follows that $p_0 = \widetilde{\mathbb{P}}\left(a_0^H, a_0^L, w_0^L\right)$ and $\underline{w} < w_0^L \leq \overline{w}$.

[22] If $p_r \cdot D(p_r) \geq p_s \cdot D(p_s)$, it can be shown that the equilibrium I find is the unique IE with $P \in \Sigma$.

Since firm $i \in \mathcal{I}^L$ makes zero profits at any $1 \leq t \leq T-1$ but wages are bounded above by $\overline{w}_1$, then output prices are bounded above by $p^{**} = \frac{r}{q(\overline{w}_1)}$. In addition, p_t solves

$$q\left[\omega^H(p_t)\right] \cdot m_t^H(h^*) + q\left[\omega^L(p_t)\right] \cdot m_t^L(h^*) = D(p_t) \qquad \text{if } 1 \leq t < T-1$$

The left side of this equation is the short run industry supply. Since $q\left[\omega^H(p_t)\right] > q\left[\omega^L(p_t)\right]$, there is an excess of demand at any price $p \leq \mathbb{P}\left(m_t^H(h^*) + m_t^L(h^*)\right)$. Assumption AD implies that for any $m_t^H(h^*) + m_t^L(h^*) < \frac{p_s \cdot D(p_s)}{r}$ there is a price $\widetilde{\mathbb{P}}\left(m_t^H(h^*), m_t^L(h^*)\right) \geq \mathbb{P}\left(m_t^H(h^*) + m_t^L(h^*)\right)$ that clears the market.

At $T-1$, the same reasoning that motivated (14) implies $m_T^H(h^*) + m_T^L(h^*) = \frac{p_s \cdot D(p_s)}{r}$ and $w_{i,T}(h^*) = \omega\left(\frac{p_s \cdot D(p_s) - r \cdot m_T^L(h^*)}{r \cdot m_{T-1}^H(h^*)}\right)$ if $i \in \mathcal{I}^H$. Since $\frac{m_T^L(h^*)}{m_{T-1}^L(h^*)} = G\left(w^L(p_{T-1})\right)$ by Proposition 4.1 (iii), p_{T-1} solves

$$q\left[\omega\left(\frac{p_s \cdot D(p_s) - r \cdot G(w^L(p_{T-1})) \cdot m_{T-1}^L(h^*)}{r \cdot m_{T-1}^H(h^*)}\right)\right] \cdot m_{T-1}^H(h^*) + q\left[\omega^L(p_{T-1})\right] \cdot m_{T-1}^L(h^*) = D(p_{T-1}) \qquad (17)$$

Clearly, $\delta = \frac{p_s \cdot D(p_s) - r \cdot G(\omega^L(p_{T-1})) \cdot m_{T-1}^L(h^*)}{r \cdot \left(p_{T-1} \cdot D(p_{T-1}) - r \cdot m_{T-1}^L(h^*)\right)}$.

Example 1: Demand is $D(p) = \frac{1}{p}$, AD holds and expenditure is always equal to 1.[23] Suppose

$$\begin{array}{ccccc} \alpha = 0.5 & \lambda \geq 0.2 & c = 1 & \overline{w}_1 = 0.2 & a_0^H = 0.14 \\ \beta = 9/10 & \frac{1}{r} = 0.9 & \theta = 0.5 & \overline{w}_2 = 8/9 & a_0^L = 0.06 \end{array}$$

I choose parameters so that AW holds. In the steady state, aggregate financial capital is $\frac{1}{r} = 0.9$, wages are $\omega(1) = 0.15$ and $v^* = 2 > \overline{w}_2$ and output price is $p_s = \frac{2}{3}\sqrt{5} > r \cdot \frac{2}{3}\sqrt{\frac{17}{5}} = p^*$. In the unique equilibrium in which all firms behave identically, the price sequence is $P = \left\{2\sqrt{6}, p_s, p_s, ...\right\}$ and $T = T_s = 1$. The entry-wage is 0.08 at date zero and 0.15 thereafter, while the probability of promotion for a young worker is 0.12 at date zero and 0.05 afterwards. Financial capital at date zero is $m_0^H(h^*) = 0.2$ and $m_t^H(h^*) = 0.9$ for all $t \geq 1$.

However, there are other equilibria in which, for example, only $2/3$ of the firms are in $\mathcal{I}^H$. I have chosen the parameters values so that in any IE either $T = 1$ or $T = 2$. This example is robust to values of μ^H around $\frac{2}{3}$. Firms in $\mathcal{I}^L$ display worst prospects for advancement than those in $\mathcal{I}^H$ iff $w_0^L > 0.08$. Hence, the date zero price must exceed $2\sqrt{6}$, the market clearing price when all firms display equal prospects for advancement. Then, $\omega^L(p_0) > \omega^L\left(2\sqrt{6}\right) > \overline{w}_1$ and so $\overline{w} = \overline{w}_1$. By (16), (17) and Proposition 4.2, $\left\{P, \frac{2}{3}, \left(\widehat{\gamma}, \widehat{\Psi}\right)\right\}$ is an IE with $T = 2$ if and only if (i) $p_0 = \widetilde{\mathbb{P}}\left(0.14, 0.06, w_0^L\right)$, (ii) $p_1 \in (p_r, p^{**}]$ and it solves (17) for $m_1^L(h^*) = G\left(w_0^L\right) \cdot a_0^L$ and $m_1^H(h^*) = p_0 \cdot q\left[\omega^H(p_0)\right] \cdot a_0^H$, (iii) $\widehat{\delta} \equiv \frac{1 - r \cdot G(\omega^L(p_1)) \cdot m_1^L(h^*)}{r \cdot (1 - r \cdot m_1^L(h^*))} \in [0, 1]$ and (iv) $p_1 \cdot q\left[\omega^H(\delta \cdot p_1)\right] > r$.

Let $u \simeq 0.105$. On the left hand side of figure 2, I plot $p_0\left(w_0^L\right) \equiv \widetilde{\mathbb{P}}\left(0.14, 0.06, w_0^L\right)$ and $p_1\left(w_0^L\right)$, the solution to (17), for each $w_0^L \in (u, 0.2]$. Set $p_0 = p_0\left(w_0^L\right)$ and $p_1 = p_1\left(w_0^L\right)$. Clearly, $P = \{p_0, p_1, p_s, p_s, ...\}$ satisfies $p_t \in [p_s, p^{**})$ for all $t \geq 1$ and then (i) and (ii) hold. The choice of u ensures $0 < m_1^H(h^*) + m_1^L(h^*) < \frac{1}{r}$ and $1 - r \cdot m_1^L(h^*) + G\left(w^L(p_1)\right) \cdot m_1^L(h^*) \geq \frac{1}{r}$. Clearly, $\delta \equiv \frac{1 - r \cdot G(\omega^L(p_1, \alpha_L)) \cdot m_1^L(h^*)}{r \cdot (1 - r \cdot m_1^L(h^*))} \leq 1$. Since $p_1 \geq p_s$ and $m_1^L(h^*) < \frac{1}{r}$, then $G\left(w^L(p_1)\right) \in (0, 1)$ and (iii) holds. Since $m_2^L(h^*) = G\left(\omega^L(p_1)\right) \cdot m_1^L(h^*) < m_1^L(h^*)$,

[23] This functional form simplifies the analysis because the industry revenue at date 1 exceeds the steady state financial capital. However, if firms follow different strategies it may take more than one period for the industry to converge to the steady state. This is because the assets of those firms that display good growth prospects may fall short of the steady state level at date 1.

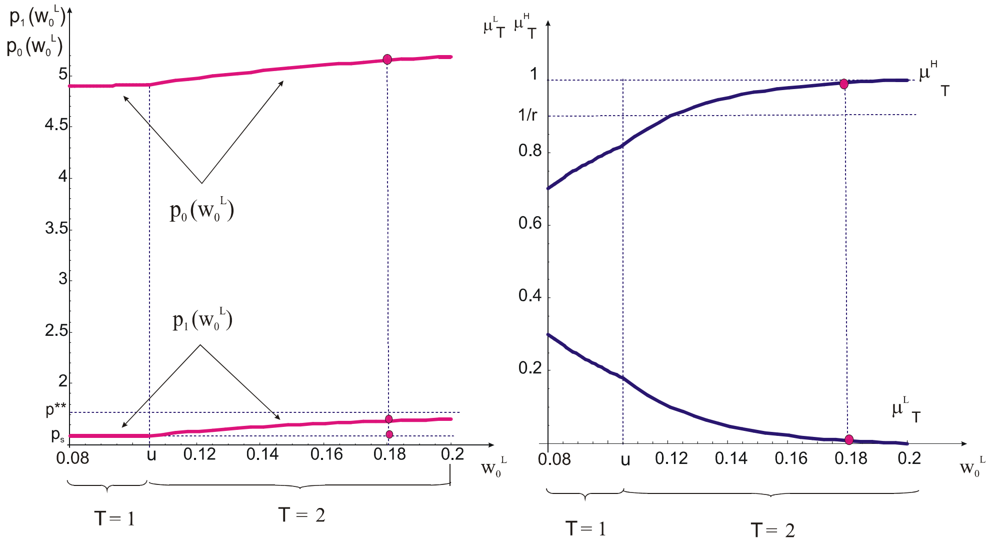

Figure 2. Equilibrium Prices and Steady State Market Shares

then $m_2^H(h^*) = \delta \cdot p_1 \cdot q\left[\omega^H(\delta \cdot p_1)\right] \cdot m_1^H(h^*) = \frac{1}{r} - m_1^L(h^*) > m_1^H(h^*)$. Since $\frac{m_2^H(h^*)}{m_1^H(h^*)} > 1$, $p_1 \cdot q\left[\omega^H(\delta \cdot p_1)\right] > p_s \cdot q[\omega(1)] = r$ and so (iv) holds. Thus, $\left\{P, \frac{2}{3}, \left(\widehat{\gamma}, \widehat{\Psi}\right)\right\}$ is an IE for any $w_0^L \in (u, 0.2]$.[24]

For each $w_0^L \in (u, 0.2]$, *ex-ante* identical firms follow different growth paths. On the right hand side of Figure 2, I plot the steady state market share of firms in $\mathcal{I}^L$ and $\mathcal{I}^H$ for each w_0^L. For large values of w_0^L, firms in $\mathcal{I}^L$ display very poor prospects for advancement at date zero and they are almost driven out at date 2. For example, if the date zero entry-wage is larger than 0.18, the steady state market share of firms in $\mathcal{I}^L$ is smaller than 1%. □

This example shows that *ex-ante* identical firms can follow different growth paths and have different sizes in the steady state. Since firms pay different wages to their young workers along the transition, technological efficiency holds only in the steady state. My analysis of the cases where firms are unconstrained or $v^* = \overline{w}_2$ implies these results depend both on financial constraints and the impossibility of enforcing a wage $\overline{w}_2$ for skilled workers.

7. Industry Equilibrium with Heterogeneous Firms

In this section, I provide the example in which firms have different technologies, one firm produces inefficiently every period but it dominates the market, in terms of market share, in the long run.

With some abuse of notation, let $q(l, s; \alpha_i)$ denote $q(l, s, 0; \alpha_i)$. In Figure 3, I illustrate the relationship between the two production functions plotting the isoquant associated with output level $\overline{q}$ for each technology.

Recall that firm $i \in \mathcal{I}^H$ if $\alpha_i = \alpha_H$ and $i \in \mathcal{I}^L$ if $\alpha_i = \alpha_L$. If both firms choose an input bundle that lies below the diagonal, firms in $\mathcal{I}^H$ produce inefficiently. Indeed, since $q(l, s; \alpha_i) = l \cdot \left(\frac{s}{l}\right)^{1-\alpha_i}$, for any $(l, s) \in \Re^2_{++}$ such that $\frac{s}{l} < 1$, it follows that $q(l, s; \alpha_L) > q(l, s; \alpha_H)$ if and only if $\alpha_H < \alpha_L$. Hence, if $\frac{s_{i,t}(h^*)}{l_{i,t}(h^*)} = \frac{1-\alpha_i}{\alpha_t} \cdot \frac{w_{i,t}(h^*)}{v^*+c} < 1$ for all $t \geq 0$ and $\alpha_H < \alpha_L$, firms in $\mathcal{I}^H$ produce inefficiently and firms in $\mathcal{I}^L$ produce efficiently every period.

[24] It can also be shown that for any $w_0^L \in (0.08, u]$ there exists an IE with $T = 1$. See Beker [5].

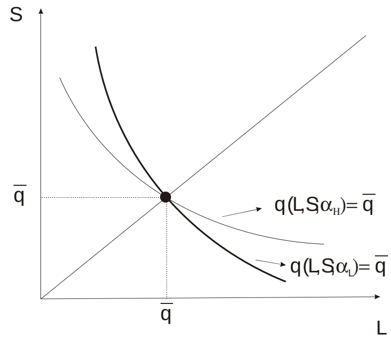

Figure 3. Isoquants of firms H and L.

In example 2, I show that if young workers are pessimistic enough about the prospects displayed by firms in $\mathcal{I}^L$, firms in $\mathcal{I}^H$ end up dominating the market. Let $w_s^i = \omega(1, \alpha_i)$ be firm i's steady state entry-wage. Since both firms must make zero profits in steady state, the technologies must satisfy:

$$q\left(w_s^H, 1\right) = q\left(w_s^L, 1\right) \Leftrightarrow \left(\frac{\alpha_H}{w_s^H}\right)^{\alpha_H} \cdot \left(\frac{1-\alpha_H}{v^*+c}\right)^{1-\alpha_H} = \left(\frac{\alpha_L}{w_s^L}\right)^{\alpha_L} \cdot \left(\frac{1-\alpha_L}{v^*+c}\right)^{1-\alpha_L} \tag{18}$$

Example 2: The demand function is $D(p) = \frac{1}{p}$, as in example 1. Suppose $\lambda = 0.95$, $\beta = \frac{1}{r} = 0.9$ and

$$\begin{array}{llll} \alpha_H = 0.1 & a_0 = 0.3 & \overline{w}_1 = \frac{50}{81} \cdot \left(\frac{125}{153}\right)^{\frac{1}{4}} \simeq 0.587 & c = 0.04 \\ \alpha_L = 0.5 & \mu^H = 0.7 & \overline{w}_2 = 0.63 & \frac{c}{\theta} = 1.33 \end{array}$$

I choose the parameters so that assumption AW and condition (18) holds. In this example, $7/10$ of the firms are in $\mathcal{I}^H$ and $3/10$ of the firms are in $\mathcal{I}^L$. The initial aggregate financial capital is 0.3 and the steady state aggregate financial capital is $\frac{1}{r} = 0.9$. Since $w_s^L \simeq 0.398$ and $w_s^H \simeq 0.113$, firms in $\mathcal{I}^H$ display better prospects for advancement than firms in $\mathcal{I}^L$ do in steady state. For any $w_0^L \in [0.276, \overline{w}_1]$, there exists an IE in which $T_s = 2$ and young workers born at date zero believe firms in $\mathcal{I}^H$ display better prospects for advancement than firms in $\mathcal{I}^L$. First, notice that there is P such that an IE with $T_s = 2$ exists only if $p_0, p_1 \in \left(p_r, \frac{r}{q(\overline{w}_1, \alpha_L)}\right]$ solve

$$\begin{aligned} q\left[\omega^H(p_0, \alpha_H), \alpha_H\right] \cdot a_0^H + q\left(w_0^L, \alpha_L\right) \cdot a_0^L &= \frac{1}{p_0} \\ q\left[\omega\left(\frac{1-r\cdot G(w^L(p_1,\alpha_L),\alpha_L)\cdot m_1^L(h^*)}{r\cdot m_1^H(h^*)}, \alpha_H\right), \alpha_H\right] \cdot m_1^H(h^*) + \frac{r}{p_1} \cdot m_1^L(h^*) &= \frac{1}{p_1} \end{aligned}$$

where $m_1^L(h^*) = G\left(w_0^L, \alpha_L\right) \cdot a_0^L$ and $m_1^H(h^*) = p_0 \cdot q\left(\omega^H(p_0, \alpha_H), \alpha_H\right) \cdot a_0^H$. I choose w_0^L such that $0 < m_1^L(h^*) + m_1^H(h^*) < \frac{1}{r}$. In the left panel of Figure 4, the dashed line corresponds to p_0, the solution to the first equation, while the full line corresponds to p_1, the solution to the second equation.

For each $w_0^L \in [0.276, \overline{w}_1]$ these equations have a unique solution that satisfies $p_0, p_1 \in \left(p_r, \frac{r}{q(\overline{w}_1, \alpha_L)}\right]$, and $w_0^L \in (\underline{w}, \overline{w})$. In addition, $1 - r \cdot m_1^L(h^*) + G\left(w^L(p_1, \alpha_L), \alpha_L\right) \cdot m_1^L(h^*) \geq \frac{1}{r}$. Clearly, $\widehat{\delta} \equiv \frac{1-r\cdot G(\omega^L(p_1,\alpha_L))\cdot m_1^L(h^*)}{r\cdot(1-r\cdot m_1^L(h^*))} \leq 1$. Since $p_1 \geq p_s(\alpha_L)$, then $G\left(w^L(p_1, \alpha_L), \alpha_L\right) \leq 1$. Since $\omega^H\left(\widehat{\delta} \cdot p_1, \alpha_H\right) = \omega\left(\frac{1-r\cdot G(w^L(p_1,\alpha_L),\alpha_L)\cdot m_1^L(h^*)}{r\cdot m_1^H(h^*)}, \alpha_H\right)$, then $\widehat{\delta} \in [0,1]$, $\frac{m_2^H(h^*)}{m_1^H(h^*)} > 1$ and $p_1 \cdot q\left(\omega\left(\frac{1-r\cdot G(w^L(p_1,\alpha_L),\alpha_L)\cdot m_1^L(h^*)}{r\cdot m_1^H(h^*)}, \alpha_H\right), \alpha_H\right) > r$. By proposition 4.2, an IE exists.

In the IE every firm makes zero profits from date 2 on. If at date zero, young workers believe firms in $\mathcal{I}^L$ display

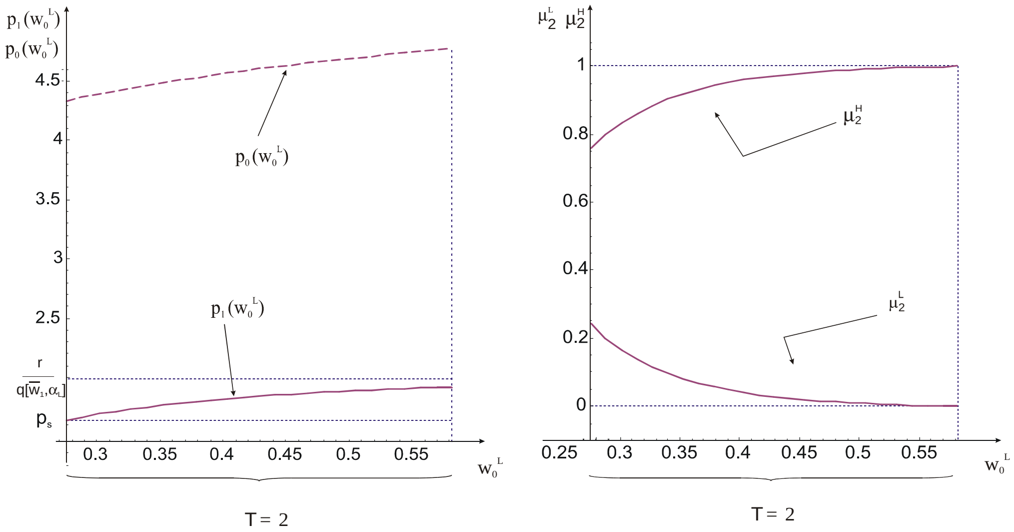

Figure 4. Equilibrium Prices and Steady State Market Shares

sufficiently bad prospects, those firms are almost driven out, in terms of market share, in the steady state. For example, if $w_0^L \geq 0.5$ then the steady state market share of firms in $\mathcal{I}^L$ is below 0.01 (see the right panel of Figure 4) which means that almost all the production is carried out by workers of firms in $\mathcal{I}^H$. In those steady states not only every firm makes zero profits, is not financially constrained and maximizes profits, but also *almost* all workers who perform the same job receive the same wage, regardless of the firm that hires them, as in a Walrasian equilibrium. However, these equilibria are productively inefficient in a strong sense. Notice that

$$\frac{s_{i,t}(h^*)}{l_{i,t}(h^*)} = \frac{1-\alpha_L}{\alpha_L} \cdot \frac{w_{i,t}(h^*)}{v^*+c} \leq \frac{\overline{w}_1}{v^*+c} < 1 \;\; \forall i \in \mathcal{I}^L \quad \text{and} \quad \frac{s_{i,t}(h^*)}{l_{i,t}(h^*)} = \frac{1-\alpha_H}{\alpha_H} \cdot \frac{w_{i,t}(h^*)}{v^*+c} \leq 9 \cdot \frac{w_s^H}{v^*+c} < 1 \;\; \forall i \in \mathcal{I}^H$$

If all labor were allocated to firm L, more output could be produced without altering workers' welfare. □

8. Conclusion

In competitive output markets, the retained earnings dynamic gives an evolutionary advantage to firms with lower unit costs. However, unit costs are determined not only by technological efficiency but also by wages. Unlike in Walrasian markets, worker's expectations about the opportunities for advancement within the firm are key to determine wages in internal labor markets. As a consequence, the fitness of a firm depends not only on its technological efficiency but also on workers' beliefs. This paper suggests that, at least in the long run, the retained earnings dynamic justifies the use of the standard static analysis of competitive markets to make positive predictions but does not justify its efficiency properties. Unlike in Blume and Easley's model [7], even the steady state of the retained earnings dynamic may fail to be efficient. In contrast with Beker [4], I do not need to assume a stochastic technology to show that inefficient firms can dominate a competitive output market. As in Arthur [2], what happens at the origin of the industry has a decisive role on the technology that dominates the market. However, it is not a network externality or the presence of increasing returns what drives the result but the young workers' beliefs about the prospects for advancement offered by the firms.

APPENDIX

A. PROOFS OF SECTIONS 3

Proof of Proposition 3.1: Suppose not. Consider $\widetilde{f}_i = \left\{\widetilde{f}_{i,t}\right\}_{t=0}^{\infty}$ where $\widetilde{f}_{i,t} : H_i^t \mapsto \Re_+^7$ is defined as follows:

a) If $x_i^\tau = h_i^\tau(h^*)$, $\widetilde{f}_{i,\tau}(x_i^\tau) = d_{i,\tau}(h)$.

b) If $t \geq \tau+1$ and x_i^t is the partial history of firm i consisting in $h_i^\tau(h^*)$ followed by the actions of the firms and workers induced by $\left\{\widetilde{f}_{i,k}, \gamma_{-i,k}\right\}_{k=\tau}^{t-1}$, consider any $h' \in H^\infty$ such that $h_i^t(h') = x_i^t$. Let $\widetilde{f}_{i,t}\left(x_i^t\right) = \Big(l_{i,t}(h^*), e_{i,t}(h^*), w_{i,t}(h^*), v_{i,t}(h^*), v_{i,t}^e(h^*), m_{i,t}(h^*), a_{i,t}(h') - m_{i,t}(h^*)\Big)$. Since $\widetilde{f}_{i,\tau}(x_i^\tau) \in \mathcal{A}_{i,\tau}(x_i^\tau)$, it suffices to show that $\widetilde{f}_{i,k}\left(x_i^k\right) \in \mathcal{A}_{i,k}\left(x_i^k\right)$ for every $\tau+1 \leq k \leq t$. Since $w_{i,k}(h^*) \cdot l_{i,k}(h^*) + v_{i,k}^e(h^*) \cdot e_{i,k}(h^*) = m_{i,k}(h^*)$, the last condition holds iff

$$a_{i,k}\left(h'\right) - m_{i,k}(h^*) > 0 \tag{19}$$

for all k such that $\tau+1 \leq k \leq t$. Since $a_{i,\tau+1}(h') - a_{i,\tau+1}(h^*) = [R_{i,\tau}(h) - R_{i,\tau}(h^*)] \cdot a_{i,\tau}(h^*) > 0$, (19) holds at $k=\tau$. So, if one proves that

$$a_{i,k}\left(h'\right) = a_{i,k}(h^*) + r^{k-1-\tau} \cdot \left[a_{i,\tau+1}\left(h'\right) - a_{i,\tau+1}(h^*)\right] \tag{20}$$

for all k such that $\tau+1 \leq k \leq t$, then (19) holds for all k such that $\tau+1 \leq k \leq t$. Clearly, (20) holds at $k=\tau+1$. We prove by induction that it holds for all $\tau+1 \leq k \leq t$. Suppose it holds up to $k \geq \tau+1$ for some $k < t$. Since $s_{i,\tau+1}(h') = s_{i,\tau+1}(h^*) \leq \lambda \cdot \widehat{l}_{i,\tau}(h) = \lambda \cdot \widehat{l}_{i,\tau}(h')$ and $s_{i,k}(h') = s_{i,k}(h^*) \leq \lambda \cdot \widehat{l}_{i,k-1}(h^*) = \lambda \cdot \widehat{l}_{i,k-1}(h')$ for any $k > \tau+1$, then $a_{i,k}(h') - a_{i,k}(h^*) = r \cdot [a_{i,k-1}(h') - a_{i,k-1}(h^*)]$ and iterating backwards one obtains (20).

c) Otherwise, let $\widetilde{f}_{i,t}\left(x_i^t\right) = f_{i,t}\left(x_i^t\right) \in \mathcal{A}_{i,t}\left(x_i^t\right)$.

The argument in (a) - (c) implies that $\widetilde{f}_i \in \mathbb{F}_i$. Let $\widetilde{h} = h^*\left(h_i^\tau(h^*), \widetilde{f}_i, \gamma\right)$. By (20), $a_{i,t}\left(\widetilde{h}\right) > a_{i,t}(h^*)$ for every $t \geq \tau+1$. A contradiction is reached because

$$\begin{aligned} E_{\Psi_{I_{i,\tau}\left(h_i^\tau(h^*)\right)}}\left[\pi_{i,\tau}\left(\widetilde{f}_i, \gamma\,|\alpha_i, P\right)\right] &= \sum_{k=\tau}^{\infty} \beta^{k+1-\tau} \cdot a_{i,k+1}\left(\widetilde{h}\right) \\ &> \sum_{k=\tau}^{\infty} \beta^{k+1-\tau} \cdot a_{i,k+1}(h^*) = E_{\Psi_{I_{i,\tau}\left(h_i^\tau(h^*)\right)}}\left(\pi_{i,\tau}\left(f_i, \gamma\,|\alpha_i, P\right)\right) \blacksquare \end{aligned}$$

Proof of Proposition 3.2: (i) Since $v_{i,t}(h^*) < \overline{w}_2$ implies $\widehat{s}_{i,t}(h^*) = 0$, then $v_{i,t}(h^*) \geq \overline{w}_2$.

(ii) Suppose $v_{i,t}(h^*) \neq v_{j,t}^e(h^*)\ \forall j \in \mathcal{D}_i$. If $v_{j,t}^e(h^*) > v_{i,t}(h^*)$ or $v_{j,t}^e(h^*) \leq \overline{w}_2\ \forall j \in \mathcal{D}_i$, choose v such that $\overline{w}_2 < v < v_{i,t}(h^*)$. Otherwise, choose v such that $\max\left\{v_{j,t}^e(h^*) : v_{j,t}^e(h^*) < v_{i,t}(h^*)\ \&\ j \in \mathcal{D}_i\right\} < v < v_{i,t}(h^*)$. Since old workers trained by firm i who stay in firm i at wage $v_{i,t}(h^*)$ still accept when offered v, then firm i has a profitable deviation. Indeed, let $d = \Big(l_{i,t}(h^*), e_{i,t}(h^*), w_{i,t}(h^*), v, v_{i,t}^e(h^*), m_{i,t}(h^*), b_{i,t}(h^*)\Big)$. Clearly, $d \in \mathcal{A}_{i,t}(h^*)$. Consider h such that

$$h^\tau(h) = h^\tau(h^*),\, d_{i,t}(h) = d,\, d_{j,\tau}(h) = d_{j,\tau}(h^*) \text{ for every } j \neq i \text{ and } h = h^*\left(\overline{h^\tau}(h), \gamma\right).$$

Since $w_{i,t}(h) = w_{i,t}(h^*)$, then $\widehat{l}_{i,t}(h) = \widehat{l}_{i,t}(h^*)$ and $\lambda \cdot \widehat{l}_{i,t}(h) \geq s_{i,t+1}(h^*)$. By Proposition 3.1, $R_{i,t}(h) \leq R_{i,t}(h^*)$.

Since $\overline{v}_{j,t}(h) = \overline{v}_{j,t}(h^*)$ for every $j \in \mathcal{S}_i$, then $e_{i,t}(h) = e_{i,t}(h^*)$. Observe that $\overline{v}_{i,t}(h) = \left[v, v_{i,t}^e(h^*), \chi_{i,t}(h^*)\right]$ and (1) in the definition of PBE implies $\widehat{e}_{i,t}(h) = \widehat{e}_{i,t}(h^*)$ and $s_{i,t}(h) \geq s_{i,t}(h^*)$. Since $\overline{v}_{j,t}(h) = \overline{v}_{j,t}(h^*)$ for every $j \in \mathcal{D}_i$, $\widehat{s}_{i,t}(h) \geq \widehat{s}_{i,t}(h^*)$ because those accepting a promotion at $v_{i,t}(h^*)$ continue to accept at v. Thus, $q_{i,t}(h) \geq q_{i,t}(h^*)$. By (9) and the assumption that $s_{i,t}(h^*) = \widehat{s}_{i,t}(h^*)$,

$$[R_{i,t}(h) - R_{i,t}(h^*)] \cdot a_{i,t}(h^*) = p_t \cdot [q_{i,t}(h) - q_{i,t}(h^*)] + c \cdot [\widehat{s}_{i,t}(h^*) - s_{i,t}(h)] + v_{i,t}(h^*) \cdot \widehat{s}_{i,t}(h^*) - v \cdot \widehat{s}_{i,t}(h).$$

Suppose $s_{i,t}(h) = \widehat{s}_{i,t}(h) = \widehat{s}_{i,t}(h^*)$. Then $[R_{i,t}(h) - R_{i,t}(h^*)] \cdot a_{i,t}(h^*) = [v_{i,t}(h^*) - v] \cdot \widehat{s}_{i,t}(h) > 0$, a contradiction. Then, either $s_{i,t}(h) > \widehat{s}_{i,t}(h)$ or $\widehat{s}_{i,t}(h) > \widehat{s}_{i,t}(h^*)$. Hence, $q_{i,t}(h) > q_{i,t}(h^*)$. Since

$$c \cdot [\widehat{s}_{i,t}(h^*) - s_{i,t}(h)] \geq v \cdot s_{i,t}(h) - v_{i,t}(h^*) \cdot \widehat{s}_{i,t}(h^*) \geq v \cdot \widehat{s}_{i,t}(h) - v_{i,t}(h^*) \cdot \widehat{s}_{i,t}(h^*),$$

a contradiction is reached because $[R_{i,t}(h) - R_{i,t}(h^*)] \cdot a_{i,t}(h^*) \geq p_t \cdot [q_{i,t}(h) - q_{i,t}(h^*)] > 0$. ■

Proof of Proposition 3.3: Suppose $\overline{w}_2 < v_{i,t}(h^*) < \frac{c}{\theta}$. Let $v^{\inf} \equiv \inf_{j \in \left\{\widetilde{j} \in \mathcal{I},\, \widehat{s}_{\widetilde{j},t}(h^*) = s_{\widetilde{j},t}(h^*)\right\}} v_{j,t}(h^*)$. Since $v^{\inf} < \frac{c}{\theta}$, there is $\varepsilon > 0$ such that $\left(v^{\inf} + \varepsilon\right) \cdot (1+\theta) < v^{\inf} + c$. Without loss in generality, suppose $v_{i,t} \leq v^{\inf} + \varepsilon$ and $v_{i,t} \leq v_{\widetilde{i},t}$ for every firm $\widetilde{i}$ that trains simultaneously with firm i, i.e. any $\widetilde{i} \in \mathcal{I}$ such that $\mathcal{D}_{\widetilde{i}} = \mathcal{D}_i$. By Proposition 3.2, $v^e_{j,t}(h^*) = v_{i,t}(h^*)$ for some $j \in \mathcal{D}_i$.

Suppose workers trained by firm i do not receive any outside offer matching firm i's offer, that is, $\widehat{v}^e_{j,t}(h^*) = 0$ for all $j \in D_i$ such that $v^e_{j,t}(h^*) = v_{i,t}(h^*)$. Firm i could profitable deviate by offering v slightly below $v_{i,t}(h^*)$ as in the proof of Proposition 3.2 (ii).

Suppose, instead, some workers trained by firm i reject the offer of some firm j that matches firm i's offer. Then, $\widehat{e}_{j,t}(h^*) < e_{j,t}(h^*)$. Let $\varepsilon' > 0$ be the number of firm i's workers who reject an firm j's offer on h^*. Since hiring workers trained by another firm for less than $\frac{c}{\theta}$ is more profitable than promoting internally, firm j could profitably deviate by offering v slightly above $v_{i,t}(h^*)$ so that every worker in firm i who receives an offer accepts. Let $d = \left(l_{j,t}(h^*), \frac{v^e_{j,t}(h^*)}{v} \cdot e_{j,t}(h^*), w_{j,t}(h^*), v_{j,t}(h^*), v, m_{j,t}(h^*), b_{j,t}(h^*)\right)$ be such that $\frac{v^e_{j,t}(h^*)}{v} \cdot e_{j,t}(h^*) > \widehat{e}_{j,t}(h^*)$ and $\varepsilon' - \left[e_{j,t}(h^*) - \frac{v^e_{j,t}(h^*)}{v} \cdot e_{j,t}(h^*)\right] > 0$. Clearly $d \in \mathcal{A}_{j,t}(h^*)$. Consider h such that

$$h^\tau(h) = h^\tau(h^*),\, d_{j,t}(h) = d,\, d_{\widetilde{j},\tau}(h) = d_{\widetilde{j},\tau}(h^*) \text{ for every } \widetilde{j} \neq j \text{ and } h = h^*\left(\overline{h^\tau}(h), \gamma\right).$$

Since $w_{j,t}(h) = w_{j,t}(h^*)$ and $l_{j,t}(h) = l_{j,t}(h^*)$, then $\widehat{l}_{j,t}(h) = \widehat{l}_{j,t}(h^*)$ and $\lambda \cdot \widehat{l}_{j,t}(h) \geq s_{j,t+1}(h^*)$. By Proposition 3.1, $R_{j,t}(h) \leq R_{j,t}(h^*)$.

Since $\overline{v}_{\widetilde{i},t}(h) = \overline{v}_{\widetilde{i},t}(h^*)$ for every $\widetilde{i} \in \mathcal{S}_j$, $\widehat{s}_{i,t}(h^*) > 0$ and $\overline{v}_{j,t}(h) = \left[v_{j,t}(h^*), v, \chi_{j,t}(h^*)\right]$, then $\widehat{e}_{j,t}(h) > \widehat{e}_{j,t}(h^*)$ because $\widehat{e}_{j,t}(h) \geq \widehat{e}_{j,t}(h^*) + \varepsilon' - [e_{j,t}(h^*) - e_{j,t}(h)]$. Since

$$s_{j,t}(h) = s_{j,t}(h^*) - \frac{v^e_{j,t}(h^*) \cdot \left[\frac{v}{v^e_{j,t}(h^*)} \cdot \widehat{e}_{j,t}(h) - \widehat{e}_{j,t}(h^*)\right]}{v_{j,t}(h^*) + c},$$

then $[R_{j,t}(h) - R_{j,t}(h^*)] \cdot a_{j,t}(h^*) \geq p_t \cdot [q_{j,t}(h) - q_{j,t}(h^*)]$. Let $\Delta = \frac{v}{v^e_{j,t}(h^*)} \cdot \widehat{e}_{j,t}(h) - \widehat{e}_{j,t}(h^*)$. Since $q_{j,t}(h) = q\left(\widehat{l}_{j,t}(h^*), \widehat{s}_{j,t}(h^*) - \frac{v^e_{j,t}(h^*) \cdot \Delta}{v_{j,t}(h^*) + c}, (\widehat{e}_{j,t}(h^*) + \Delta) \cdot \frac{v^e_{j,t}(h^*)}{v}, \alpha_j\right)$ and $q\left(l, s' - \frac{v^e_{j,t}(h^*) \cdot \Delta}{v_{j,t}(h^*) + c}, e + \Delta, \alpha_j\right)$ is strictly increasing in Δ, it follows that $[q_{j,t}(h) - q_{j,t}(h^*)] > 0$ for v close to $v^e_{j,t}(h^*)$. Indeed,

$$\frac{d}{d\Delta} q\left(l, s' - \frac{v^e_{j,t}(h^*) \cdot \Delta}{v_{j,t}(h^*) + c}, e + \Delta, \alpha_j\right) = \frac{\partial}{\partial s} q\left(l, s' - \frac{v^e_{j,t}(h^*) \cdot \Delta}{v_{j,t}(h^*) + c}, e + \Delta, \alpha_j\right) \cdot \left(\frac{1}{1+\theta} - \frac{v^e_{j,t}(h^*)}{v_{j,t}(h^*) + c}\right) > 0$$

where I use the fact that $v^e_{j,t}(h^*) = v_{i,t}(h^*)$ and $\frac{1}{1+\theta} > \frac{v_{i,t}(h^*)}{v_{j,t}(h^*) + c}$. Hence $R_{j,t}(h) > R_{j,t}(h^*)$, a contradiction. ■

Proof of Proposition 3.4: Since $\widehat{l}_{i,t}(h^*) > 0$ and by (ii) in the definition of PBE, $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}[U_{i,t}(\sigma_{i,t},\gamma)] \geq \overline{w}_1 + \beta \cdot \overline{w}_2$. Suppose $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}[U_{i,t}(\sigma_{i,t},\gamma)] > \overline{w}_1 + \beta \cdot \overline{w}_2$. By hypothesis, there is $w < w_{i,t}(h^*)$ such that $E_{\Psi_{I^1_{i,t}(w)}}[U_{i,t}(\sigma_{i,t},\gamma)] > \overline{w}_1 + \beta \cdot \overline{w}_2$. If $\chi_{i,t}(h^*) = 0$, choose w so close to $w_{i,t}(h^*)$ that $\frac{m_{i,t}(h^*) - w \cdot l_{i,t}(h^*)}{v_{i,t}(h^*)+c} < \lambda \cdot l_{i,t-1}(h^*)$. Condition (3) in the definition of PBE implies that young workers accept employment on $I^1_{i,t}(w)$. Let $\widetilde{m} = m_{i,t}(h^*) - [w_{i,t}(h^*) - w] \cdot l_{i,t}(h^*)$, $\widetilde{b} = a_{i,t}(h^*) - \widetilde{m}$ and $d = \left(l_{i,t}(h^*), e_{i,t}(h^*), w, v_{i,t}(h^*), v^e_{i,t}(h^*), \widetilde{m}, \widetilde{b}\right)$. Clearly $d \in \mathcal{A}_{i,t}(h^*)$. Consider h such that

$$h^{\tau}(h) = h^{\tau}(h^*), \, d_{i,t}(h) = d \text{ and } h = h^*\left(\overline{h^{\tau}}(h), \gamma\right) \text{ and } d_{j,\tau}(h) = d_{j,\tau}(h^*) \text{ for every } j \neq i.$$

It is easy to see that $\lambda \cdot \widehat{l}_{i,t}(h) = \lambda \cdot \widehat{l}_{i,t}(h^*) \geq s_{i,t+1}(h^*)$ but $[R_{i,t}(h) - R_{i,t}(h^*)] \cdot a_{i,t}(h^*) = r \cdot [w_{i,t}(h^*) - w] \cdot l_{i,t}(h^*) > 0$ which contradicts Proposition 3.1.■

Proof of Lemma 3.1: By definition, firm i displays better prospects than firm j if $w\left(\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)}, v_{i,t+1}(h^*)\right) < w\left(\frac{s_{j,t+1}(h^*)}{\widehat{l}_{j,t}(h^*)}, v_{j,t+1}(h^*)\right)$. Since $v_{i,t+1}(h^*) = v_{j,t+1}(h^*)$, this holds if and only if $\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)} > \frac{s_{j,t+1}(h^*)}{\widehat{l}_{j,t}(h^*)}$. ■

Proof of Proposition 3.5: Suppose $\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)} = \lambda$. Since $w_{i,t}(h^*) \geq 0$, $v_{i,t+1}(h^*) \geq v^* > \overline{w}_2$ and AW holds, $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}[U_{i,t}(\sigma_{i,t},\gamma)] \geq \beta \cdot (\lambda \cdot [v^* - \overline{w}_2] + \overline{w}_2) > \overline{w}_1 + \beta \cdot \overline{w}_2$, a contradiction. ■

Proof of Proposition 3.6: (i) Since $E_{\Psi_{I^1_{i,t}(w_{i,t}(h^*))}}[U_{i,t}(\sigma_{i,t},\gamma)] = w_{i,t}(h^*) + \beta \cdot \frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)} \cdot (v^* - \overline{w}_2)$ and $w_{i,t}(h^*) = w\left(\frac{s_{i,t+1}(h^*)}{\widehat{l}_{i,t}(h^*)}, v^*\right)$.

(ii) Suppose not. Then, $s_{i,t}(h^*) > \widehat{s}_{i,t}(h^*)$ for some i. Let $e' \equiv \widehat{s}_{i,t}(h^*) \cdot (1+\theta) + \widehat{e}_{i,t}(h^*)$ be the number of external trained workers firm i would need to produce $q_{i,t}(h^*)$ if it hired $\widehat{l}_{i,t}(h^*)$ young worker and it did not employ internally promoted workers. Let $s' \equiv \max\left\{0, \frac{e'}{1+\theta} - \sum_{j \in \mathcal{S}_i} \frac{s_{j,t}(h^*)}{1+\theta}\right\}$ be the internal promotions firm i would need to produce $q_{i,t}(h^*)$ if it hired $\widehat{l}_{i,t}(h^*)$ but faced a supply of $\sum_{j \in \mathcal{S}_i} s_{j,t}(h^*)$ externally trained workers. Let $\widetilde{m} \equiv v \cdot e' + w_{i,t}(h^*) \cdot l_{i,t}(h^*)$ and $\widetilde{b} \equiv a_{i,t}(h^*) - \widetilde{m}$. Using the assumption that $v^* = \frac{c}{\theta}$, after some algebra, one obtains that $\widetilde{m} = m_{i,t}(h^*) - [(v^* - v) \cdot e' + (v^* + c) \cdot (s_{i,t}(h^*) - \widehat{s}_{i,t}(h^*))]$. Since $s_{i,t}(h^*) > \widehat{s}_{i,t}(h^*)$, one can choose $v > v^*$ such that $0 < (v^* - v) \cdot e' + c \cdot [s_{i,t}(h^*) - \widehat{s}_{i,t}(h^*)] < m_{i,t}(h^*)$ and so $0 < \widetilde{m} < m_{i,t}(h^*)$. Let $d \equiv \left[l_{i,t}(h^*), e', w_{i,t}(h^*), v, v, \widetilde{m}, \widetilde{b}\right]$. That $d \in \mathcal{A}_{i,t}(h^*)$ follows by definition of $\widetilde{m}$ and $\widetilde{b}$ and because $\widetilde{m} < m_{i,t}(h^*)$. Then, there is h such that

$$h^t(h) = h^t(h^*), \, d_{i,t}(h) = d, \, d_{j,t}(h) = d_{j,t}(h^*) \text{ for every } j \neq i \text{ and } h = h^*\left(\overline{h^t}(h), \gamma\right).$$

Since $w_{i,t}(h) = w_{i,t}(h^*)$, then $\widehat{l}_{i,t}(h) = \widehat{l}_{i,t}(h^*)$ and $\lambda \cdot \widehat{l}_{i,t}(h) \geq s_{i,t+1}(h^*)$. By Proposition 3.1, $R_{i,t}(h) \leq R_{i,t}(h^*)$.

Since $v^e_{i,t}(h) > v_{j,t}(h)$ for all $j \in \mathcal{S}_i$, $\widehat{e}_{i,t}(h) = \min\left\{e', \sum_{j \in \mathcal{S}_i} s_{j,t}(h^*)\right\}$. Since $v_{i,t}(h) > v^e_{j,t}(h)$ for all $j \in \mathcal{D}_i$, $s_{i,t}(h) = \widehat{s}_{i,t}(h)$. Moreover, since $s' \leq \widehat{s}_{i,t}(h^*) < \lambda \cdot \widehat{l}_{i,t-1}(h^*)$, it follows that $s_{i,t}(h) = \widehat{s}_{i,t}(h) = s'$. Then, $\widehat{s}_{i,t}(h) + \frac{\widehat{e}_{i,t}(h)}{1+\theta} = \frac{e'}{1+\theta} = \widehat{s}_{i,t}(h^*) + \frac{\widehat{e}_{i,t}(h^*)}{1+\theta}$ and, therefore, $q_{i,t}(h) = q_{i,t}(h^*)$. Consequently,

$$m_{i,t}(h^*) - \left[v^e_{i,t}(h^*) \cdot \widehat{e}_{i,t}(h^*) + c \cdot s_{i,t}(h^*) + v_{i,t}(h^*) \cdot \widehat{s}_{i,t}(h^*)\right] = v^* \cdot [s_{i,t}(h^*) - \widehat{s}_{i,t}(h^*)] \quad (21)$$

$$m_{i,t}(h) - \left[v^e_{i,t}(h) \cdot \widehat{e}_{i,t}(h) + c \cdot s_{i,t}(h) + v_{i,t}(h) \cdot \widehat{s}_{i,t}(h)\right] = 0 \quad (22)$$

Finally, (21), (22) and $q_{i,t}(h) = q_{i,t}(h^*)$ implies that

$$\begin{aligned}[R_{i,t}(h) - R_{i,t}(h^*)] \cdot a_{i,t}(h^*) &= r \cdot [b_{i,t}(h) - b_{i,t}(h^*)] - v^* \cdot [s_{i,t}(h^*) - \widehat{s}_{i,t}(h^*)] \\ &= [r \cdot (v^* + c) - v^*] \cdot [s_{i,t}(h^*) - \widehat{s}_{i,t}(h^*)] + r \cdot (v^* - v) \cdot e' \\ &> c \cdot [s_{i,t}(h^*) - \widehat{s}_{i,t}(h^*)] + (v^* - v) \cdot (s' + e') > 0\end{aligned}$$

where the first inequality follows because $r > 1$. But this contradicts $R_{i,t}(h) \leq R_{i,t}(h^*)$.

(iii) If $\left[\widehat{l}_{i,t}(h^*), s_{i,t}(h^*)\right]$ does not solve (10), let $d_\varepsilon = \left(l_\varepsilon, e_\varepsilon, w_{i,t}(h^*), v_{i,t}(h^*), v^e_{i,t}(h^*), m_{i,t}(h^*), b_{i,t}(h^*)\right)$ where $(l_\varepsilon, e_\varepsilon) = (1-\varepsilon) \cdot \left(l_{i,t}(h^*), \frac{v^*+c}{v^*} \cdot s_{i,t}(h^*)\right) + \varepsilon \cdot \left(\frac{\alpha_i \cdot m_{i,t}(h^*)}{w_{i,t}(h^*)}, \frac{(1-\alpha_i) \cdot m_{i,t}(h^*)}{v^*}\right)$ and $\varepsilon \in (0,1)$. Clearly, $d_\varepsilon \in \mathcal{A}_{i,t}(h^*)$. By the hypothesis and the definition of $e_{\widetilde{\varepsilon}}$, there is $\widetilde{\varepsilon} \in (0,1)$ such that $\frac{v^*}{v^*+c} \cdot e_{\widetilde{\varepsilon}} < \lambda \cdot \widehat{l}_{i,t-1}(h^*)$ and $s_{i,t+1}(h^*) < \lambda \cdot \left[(1-\widetilde{\varepsilon}) \cdot l_{i,t}(h^*) + \widetilde{\varepsilon} \cdot \frac{\alpha_i \cdot m_{i,t}(h^*)}{w_{i,t}(h^*)}\right]$. Consider h such that

$$h^t(h) = h^t(h^*), d_{i,t}(h) = d_{\widetilde{\varepsilon}}, d_{j,t}(h) = d_{j,t}(h^*) \text{ for every } j \neq i \text{ and } h = h^*\left(\overline{h^t}(h), \gamma\right).$$

Since $w_{i,t}(h) = w_{i,t}(h^*)$, then $\widehat{l}_{i,t}(h) = l_{i,t}(h) = l_{\widetilde{\varepsilon}}$ and $\lambda \cdot \widehat{l}_{i,t}(h) > s_{i,t+1}(h^*)$. By Proposition 3.1, $R_{i,t}(h) \leq R_{i,t}(h^*)$. Since $\widehat{e}_{i,t}(h^*) = 0$, then $\frac{m_{i,t}(h^*) - w_{i,t}(h^*) \cdot l_{i,t}(h^*)}{v^*+c} = s_{i,t}(h^*) < \lambda \cdot \widehat{l}_{i,t-1}(h^*)$. Also, $\frac{m_{i,t}(h) - w_{i,t}(h) \cdot l_{i,t}(h)}{v^*+c} = \frac{v^*}{v^*+c} \cdot e_{\widetilde{\varepsilon}} < \lambda \cdot \widehat{l}_{i,t-1}(h^*)$. So $\chi_{i,t}(h) = \chi_{i,t}(h^*) = 0$. Since $\overline{v}_{j,t}(h) = \overline{v}_{j,t}(h^*)$ for every $j \in \mathcal{I}$, then $\widehat{e}_{j,t}(h) = \widehat{e}_{j,t}(h^*) = 0$ for every $j \in \mathcal{I}$ and $\widehat{s}_{i,t}(h) = s_{i,t}(h) = \frac{v^*}{v^*+c} \cdot e_{\widetilde{\varepsilon}}$. Let $R^*(p, w, \alpha) \equiv p \cdot \left(\frac{\alpha}{w}\right)^\alpha \cdot \left(\frac{1-\alpha}{v^*+c}\right)^{1-\alpha}$. It follows that $[R_{i,t}(h) - R_{i,t}(h^*)] = \widetilde{\varepsilon} \cdot [R^*(p_t, w_{i,t}(h^*), \alpha_i) - R_{i,t}(h^*)] > 0$, a contradiction.

(iv) It follows trivially from (i) and (ii). ■

Proof of Lemma 3.3 If $\frac{c}{\theta} = \overline{w}_2$, $\omega^H(\cdot, \alpha) = \overline{w}_1$. Suppose $\frac{c}{\theta} > \overline{w}_2$. Let $p \geq 0$, $\alpha \in (0,1)$ and $B(w,p) \equiv \overline{w}_1 - \beta \cdot \frac{1-\alpha}{\alpha} \cdot \frac{w}{v^*+c} \cdot p \cdot q^i(w, \alpha) \cdot (v^* - \overline{w}_2) - w$. A solution to (12) exists iff there is w such that $B(w,p) = 0$. Notice that $B(\overline{w}_1, p) \leq 0$ and $\underset{w \to 0}{Lim} B(w,p) = \overline{w}_1 > 0$. Since B is continuous and strictly decreasing in w, there is a unique solution $w^H(p, \alpha)$ to (12). Let $0 < p_1 < p_2$. Since $B\left(w^H(p_2, \alpha), p_2\right) = 0 = B\left(w^H(p_1, \alpha), p_1\right) > B\left(w^H(p_1, \alpha), p_2\right)$ then $w^H(p_2, \alpha) < w^H(p_1, \alpha)$. Hence, w^H is strictly decreasing in p. ■

Lemma A.1 *Suppose AW holds, $P \in \Sigma$ and (γ, Ψ) is a *PBESO. If $m_{i,t}(h^*) < p_{t-1} \cdot q(w_{i,t-1}(h^*), \alpha_i) \cdot m_{i,t-1}(h^*)$ then $p_t \cdot q(w_{i,t}(h^*), \alpha_i) \leq r$.*

Proof of Lemma A.1: Suppose $p_t \cdot q(w_{i,t}(h^*), \alpha_i) > r$. Define $m_\varepsilon = \varepsilon \cdot a_{i,t}(h^*) + (1-\varepsilon) \cdot m_{i,t}(h^*)$, $b_\varepsilon = a_{i,t}(h^*) - m_\varepsilon$, $l_\varepsilon = \frac{1-\alpha_i}{w_{i,t}(h^*)} \cdot m_\varepsilon$, $e_\varepsilon = \frac{1-\alpha_i}{v^*} \cdot m_\varepsilon$ and $d_\varepsilon = (l_\varepsilon, e_\varepsilon, w_{i,t}(h^*), v^*, v^*, m_\varepsilon, b_\varepsilon)$. By Propositions 3.5 and 3.6 (ii), there is $\widetilde{\varepsilon}$ such that $\lambda \cdot l_{\widetilde{\varepsilon}} > s_{i,t+1}(h^*)$ and $\lambda \cdot l_{i,t-1}(h^*) > \frac{1-\alpha_i}{v^*+c} \cdot m_\varepsilon$. Clearly, $d_{\widetilde{\varepsilon}} \in \mathcal{A}_{i,t}(h^*)$. Consider h such that $h^\tau(h) = h^\tau(h^*)$, $d_{i,t}(h) = d_{\widetilde{\varepsilon}}$, $d_{j,\tau}(h) = d_{j,\tau}(h^*)$ for all $j \neq i$ and $h = h^*\left(\overline{h^\tau}(h), \gamma\right)$. Since $w_{i,t}(h) = w_{i,t}(h^*)$, then $\widehat{l}_{i,t}(h) = l_{\widetilde{\varepsilon}}$ and $\lambda \cdot \widehat{l}_{i,t}(h) > s_{i,t+1}(h^*)$. By Proposition 3.1, $R_{i,t}(h) \leq R_{i,t}(h^*)$. A reasoning analogous to the one used in Proposition 3.6 (ii) shows that $\widehat{e}_{i,t}(h) = 0$ and $\widehat{s}_{i,t}(h) = \widehat{s}_{i,t}(h^*)$. Then, $[R_{i,t}(h) - R_{i,t}(h^*)] \cdot a_{i,t}(h^*) = [p_t \cdot q(w_{i,t}(h^*), \alpha_i) - r] \cdot \widetilde{\varepsilon} \cdot [a_{i,t}(h^*) - m_{i,t}(h^*)] > 0$, a contradiction. ■

Proof of Proposition 3.7: First I show that $p_t \cdot q(w_{i,t}(h^*), \alpha_i) \leq r \Rightarrow p_{t+1} \cdot q(w_{i,t+1}(h^*), \alpha_i) \leq r$. Since $p_t > p_r$, then $w_{i,t}(h^*) \geq w^L(p_t, \alpha_i) > \omega^H(p_t, \alpha_i)$. Since $w_{i,t}(h^*) = \omega\left(\frac{m_{i,t+1}(h^*)}{m_{i,t}(h^*)}, \alpha_i\right)$ by Proposition 3.6, then $\frac{m_{i,t+1}(h^*)}{m_{i,t}(h^*)} < p_t \cdot q(w_{i,t}(h^*), \alpha_i) \leq r$. By Lemma A.1, $p_{t+1} \cdot q(w_{i,t+1}(h^*), \alpha_i) \leq r$.

From this, it follows trivially that $p_\tau \cdot q\left(w_{i,\tau}\left(h^*\right), \alpha_i\right) = r \Rightarrow p_t \cdot q\left(w_{i,t}\left(h^*\right), \alpha_i\right) \leq r$ for every $t \geq \tau + 1$.

Let τ be the first t such that $w_{i,t}\left(h^*\right) \neq \omega^H\left(p_t, \alpha_i\right)$. Since $p_t > p_r$ and $w_{i,t}\left(h^*\right) = \omega^H\left(p_t, \alpha_i\right)$ for all $t < \tau$, then $p_t \cdot q\left(w_{i,t}\left(h^*\right), \alpha_i\right) > r$ for all $t < \tau$. By Lemma A.1, $m_{i,t}\left(h^*\right) = a_{i,t}\left(h^*\right)$ for all $t < \tau$. Since $m_{i,\tau}\left(h^*\right) = G\left(w_{i,\tau-1}\left(h^*\right), \alpha_i\right) \cdot a_{i,\tau-1}\left(h^*\right)$, then $m_{i,\tau}\left(h^*\right) = a_{i,\tau}\left(h^*\right)$. Since $m_{i,\tau+1}\left(h^*\right) \leq a_{i,\tau+1}\left(h^*\right) = p_\tau \cdot q\left(w_{i,\tau}\left(h^*\right), \alpha_i\right) \cdot a_{i,\tau}\left(h^*\right)$ and $w_{i,\tau}\left(h^*\right) \neq \omega^H\left(p_\tau, \alpha_i\right)$, then $\frac{m_{i,\tau+1}\left(h^*\right)}{m_{i,\tau}\left(h^*\right)} < p_\tau \cdot q\left(w_{i,\tau}\left(h^*\right), \alpha_i\right)$. Hence, $w_{i,\tau}\left(h^*\right) \in \left(\omega^H\left(p_\tau, \alpha_i\right), \overline{w}_1\right]$. By Lemma A.1, $m_{i,\tau+1}\left(h^*\right) < p_\tau \cdot q\left(w_{i,\tau}\left(h^*\right), \alpha_i\right) \cdot m_{i,\tau}\left(h^*\right)$ implies that $p_{\tau+1} \cdot q\left(w_{i,\tau+1}\left(h^*\right), \alpha_i\right) \leq r$. By the argument above, $p_t \cdot q\left(w_{i,t}\left(h^*\right), \alpha_i\right) \leq r$ for all $t \geq \tau + 1$. Suppose there is $t \geq \tau + 1$ such that $p_t \cdot q\left(w_{i,t}\left(h^*\right), \alpha_i\right) < r$ and $m_{i,t}\left(h^*\right) > 0$. Clearly, $R_{i,t}\left(h^*\right) < r$ and $R_{i,k}\left(h^*\right) \leq r$ for any $k \geq \tau + 1$. Consider $\widetilde{f}_i = \left\{\widetilde{f}_{i,k}\right\}_{t=0}^{\infty}$ where $\widetilde{f}_{i,k} : H_i^k \mapsto \mathcal{A}_{i,k}$ is defined as follows. If $k < \tau$ or $\tau \leq k$ and $h_i^\tau\left(h\right) \neq h_i^\tau\left(h^*\right)$, then $\widetilde{f}_{i,k}\left(h_i^k\left(h\right)\right) = f_{i,k}\left(h_i^k\left(h\right)\right)$. If $k \geq \tau$ and $h_i^\tau\left(h\right) = h_i^\tau\left(h^*\right)$, then $\widetilde{f}_{i,k}\left(h_i^k\left(h\right)\right)$ is such that $b_{i,k}\left(h_i^k\left(h\right)\right) = a_{i,k}\left(h\right)$. Hence, $\widetilde{f}_i \in \mathbb{F}_i$ and

$$\begin{aligned} E_{\Psi_{I_{i,t}\left(h_i^t\left(h^*\right)\right)}}\left[\pi_{i,t}\left(\widetilde{f}_i, \gamma\left|\alpha_i, P\right.\right)\right] &= \tfrac{\beta\cdot r}{1-\beta\cdot r}\cdot a_{i,t}\left(h^*\right) \\ &> \left[\beta\cdot R_{i,t}\left(h^*\right) + \tfrac{\beta^2\cdot r}{1-\beta\cdot r}\right]\cdot a_{i,t}\left(h^*\right) \geq E_{\Psi_{I_{i,\tau}\left(h_i^t\left(h^*\right)\right)}}\left[\pi_{i,t}\left(f_i, \gamma\left|\alpha_i, P\right.\right)\right] \end{aligned}$$

a contradiction. One concludes that $p_t \cdot q\left(w_{i,t}\left(h^*\right), \alpha_i\right) = r$ for every $t \geq \tau + 1$ such that $m_{i,t}\left(h^*\right) > 0$. ■

B. Proofs of Section 4

I begin this section with a lemma that puts a bound on the growth rate of financial capital.

Lemma A.2 *Suppose AW holds and $h' = h^*\left(x^t, \widehat{\gamma}\right)$ for some $x^t \in H^t$. Then,* $\frac{m_{i,t}\left(h'\right)}{m_{i,t-1}\left(h'\right)} \leq G\left(w_{i,t-1}\left(h'\right), \alpha_i\right)$.

Proof of Lemma A.2: Suppose $i \in \mathcal{I}^L$. If $t = 1$, then $m_{i,t}\left(h'\right) \leq G\left(w_0^L, \alpha_i\right) \cdot m_{i,t-1}\left(h'\right) = G\left(w_{i,t-1}\left(h'\right), \alpha_i\right) \cdot m_{i,t-1}\left(h'\right)$. If $t \geq 2$, $m_{i,t}\left(h'\right) \leq G\left(\omega^L\left(p_t, \alpha_i\right), \alpha_i\right) \cdot m_{i,t-1}\left(h'\right) = G\left(w_{i,t-1}\left(h'\right), \alpha_i\right) \cdot m_{i,t-1}\left(h'\right)$.

Suppose $i \in \mathcal{I}^H$. Let $\delta_t = 1$ if $t < T - 1$ and $\delta_{T-1} = \delta$. If $t \leq T - 1$, $m_{i,t}\left(h'\right) = a_{i,t}\left(h'\right)$ and

$$\begin{aligned} m_{i,t+1}\left(h'\right) = a_{i,t+1}\left(h'\right) &\leq \delta_t \cdot p_t \cdot \left(\tfrac{\alpha_i}{w_{i,t}\left(h'\right)}\right)^{\alpha_i}\left(\tfrac{1-\alpha_i}{v^*+c}\right)^{1-\alpha_i} a_{i,t}\left(h'\right) \\ &= \delta_t \cdot p_t \cdot \left(\tfrac{\alpha_i}{\omega^H\left(\delta_t\cdot p_t, \alpha_i\right)}\right)^{\alpha_i}\left(\tfrac{1-\alpha_i}{v^*+c}\right)^{1-\alpha_i} a_{i,t}\left(h'\right) \\ &= G\left[\omega^H\left(\delta_t \cdot p_t, \alpha_i\right), \alpha_i\right] \cdot a_{i,t}\left(h'\right) = G\left(w_{i,t}\left(h'\right), \alpha_i\right) \cdot m_{i,t}\left(h'\right) \end{aligned}$$

If $t \geq T$, then $m_{i,t+1}\left(h'\right) \leq m_{i,t}\left(h'\right) = G\left[\omega\left(1, \alpha_i\right), \alpha_i\right] \cdot m_{i,t}\left(h'\right) = G\left(w_{i,t}\left(h'\right), \alpha_i\right) \cdot m_{i,t}\left(h'\right)$. It follows that $m_{i,t+1}\left(h'\right) \leq G\left(w_{i,t}\left(h'\right), \alpha_i\right) \cdot m_{i,t}\left(h'\right)$ for every $t \geq 0$ and $i \in \mathcal{I}$. ■

Proof of Proposition 4.1: First notice that $v_{i,t}\left(\widehat{h}\right) = v_{i,t}^e\left(\widehat{h}\right) = v^*$ for every $t \geq 0$ and $i \in \mathcal{I}$. Since $w_{i,t}\left(\widehat{h}\right)$ equals the cutoff value of the young worker's strategy at t, then $\widehat{l}_{i,t}\left(\widehat{h}\right) = \widetilde{l}_t\left(\alpha_i\right)\left(l_{i,t-1}\left(\widehat{h}\right), m_{i,t-1}\left(\widehat{h}\right), a_{i,t}\left(\widehat{h}\right)\right) = \frac{\alpha_i}{w_{i,t}\left(\widehat{h}\right)} \cdot m_{i,t}\left(\widehat{h}\right)$ and so (i) holds at every date $t \geq 0$ and $i \in \mathcal{I}$.

Consider date 0. Since $\frac{a_{i,0}\left(\widehat{h}\right) - w_{i,0}\left(\widehat{h}\right)\cdot l_{i,0}\left(\widehat{h}\right)}{v^*+c} = \frac{1-\alpha_i}{v^*+c}\cdot a_0$ and $\frac{1-\alpha_i}{v^*+c}\cdot a_0 < \lambda\cdot l_{-1} = \lambda\cdot\widehat{l}_{i,-1}\left(\widehat{h}\right)$, then $\chi_{i,0}\left(\widehat{h}\right) = 0$ and so workers trained by firm $j \in \mathcal{S}_i$ do not join firm i, that is $\widehat{e}_{i,0}\left(\widehat{h}\right) = 0$. Since $\frac{a_0 - w_{i,0}\left(\widehat{h}\right)\cdot l_{i,0}\left(\widehat{h}\right)}{v^*+c} = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,0}\left(\widehat{h}\right) < \lambda \cdot \widehat{l}_{i,-1}\left(\widehat{h}\right)$, then $s_{i,0}\left(\widehat{h}\right) = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,0}\left(\widehat{h}\right) < \lambda \cdot \widehat{l}_{i,-1}\left(\widehat{h}\right)$. Hence, (ii) holds at date 0. Finally, notice that $m_{i,0}\left(\widehat{h}\right) = \widetilde{m}_0\left(T, \delta\right)\left(l_{-1}, 0, a_0\right) = a_0$ if $i \in \mathcal{I}^H$ and $m_{i,0}\left(\widehat{h}\right) = \widetilde{m}_0\left(w_0^L\left|\alpha_i, P\right.\right)\left(l_{-1}, 0, a_0\right) = a_0$ if $i \in \mathcal{I}^L$. Thus, (iii) holds at 0.

Now suppose (i) - (iii) holds at every date $\tau \leq t-1$ for some $t \geq 1$. Since $l_{i,t}\left(\widehat{h}\right) = \frac{\alpha_i}{w_{i,t}\left(\widehat{h}\right)} \cdot m_{i,t}\left(\widehat{h}\right)$, it follows that $\frac{m_{i,t}\left(\widehat{h}\right) - w_{i,t}\left(\widehat{h}\right) \cdot l_{i,t}\left(\widehat{h}\right)}{v^*+c} = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,t}\left(\widehat{h}\right)$. Then, to show that (ii) holds, it suffices to show that $\chi_{i,t}\left(\widehat{h}\right) = 0$. Since $\widehat{l}_{i,t-1}\left(\widehat{h}\right) = \frac{\alpha_i}{w_{i,t-1}\left(\widehat{h}\right)} \cdot m_{i,t-1}\left(\widehat{h}\right)$, it suffices to show that $\frac{1-\alpha_i}{\alpha_i} \cdot \frac{w_{i,t-1}\left(\widehat{h}\right)}{v^*+c} \cdot \frac{m_{i,t}\left(\widehat{h}\right)}{m_{i,t-1}\left(\widehat{h}\right)} < \lambda$. If AW holds, then using Lemma A.2 and that $\frac{1-\alpha_i}{\alpha_i} \cdot \frac{w}{v^*+c} \cdot G\left(w, \alpha_i\right) < \lambda$ for any $w \geq 0$ one concludes $\chi_{i,t}\left(\widehat{h}\right) = 0$. If $\frac{c}{\theta} = \overline{w}_2$, $p_0 \leq \frac{p^*(\alpha_i)}{r} \cdot \frac{\alpha_i}{1-\alpha_i} \cdot \frac{\overline{w}_2+c}{\overline{w}_1} \cdot \lambda$ and P is non-increasing, then $\frac{1-\alpha_i}{\alpha_i} \cdot \frac{w_{i,t-1}\left(\widehat{h}\right)}{v^*+c} \cdot \frac{m_{i,t}\left(\widehat{h}\right)}{m_{i,t-1}\left(\widehat{h}\right)} \leq \frac{1-\alpha_i}{\alpha_i} \cdot \frac{w_{i,t-1}\left(\widehat{h}\right)}{v^*+c} \cdot p_0 \cdot q\left(\overline{w}_1, \alpha_i\right) \leq \frac{1-\alpha_i}{\alpha_i} \cdot \frac{w_{i,t-1}\left(\widehat{h}\right)}{v^*+c} \cdot p_0 \cdot \frac{r}{p^*(\alpha_i)} < \lambda$. Therefore, for every $j \in \mathcal{S}_i$ workers trained by firm j prefer to stay in j rather than moving to i. Hence, $\widehat{e}_{i,t}\left(\widehat{h}\right) = 0$ and $s_{i,t}\left(\widehat{h}\right) = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,t}\left(\widehat{h}\right) < \lambda \cdot \widehat{l}_{i,-1}\left(\widehat{h}\right)$ so that (ii) holds at date t.

To show that (iii) holds at t notice that since $\widehat{e}_{j,t-1}\left(\widehat{h}\right) = 0$ and $v_{i,t-1}\left(\widehat{h}\right) = v^* \geq \overline{w}_2$, workers trained by firm i at $t-1$ did not switch firms, i.e. $\widehat{s}_{i,t-1}\left(\widehat{h}\right) = s_{i,t-1}\left(\widehat{h}\right)$. Then,

$$R_{i,t-1}\left(\widehat{h}\right) = R^*\left(p_{t-1}, w_{i,t-1}\left(\widehat{h}\right), \alpha_i\right) \cdot \frac{m_{i,t-1}\left(\widehat{h}\right)}{a_{i,t-1}\left(\widehat{h}\right)} + r \cdot \frac{b_{i,t-1}\left(\widehat{h}\right)}{a_{i,t-1}\left(\widehat{h}\right)} \geq r > 1$$

and so $a_{i,t}\left(\widehat{h}\right) > a_{i,t-1}\left(\widehat{h}\right)$. Consider $i \in \mathcal{I}^H$ and recall that $m_{i,t}\left(\widehat{h}\right) = \widetilde{m}_t\left(T, \delta\right)\left[l_{i,t-1}\left(\widehat{h}\right), m_{i,t-1}\left(\widehat{h}\right), a_{i,t}\left(\widehat{h}\right)\right]$. If $t < T$, then $m_{i,t}\left(\widehat{h}\right) = a_{i,t}\left(\widehat{h}\right)$. If $t \geq T$, then $a_{i,t}\left(\widehat{h}\right) > a_{i,t-1}\left(\widehat{h}\right) \geq m_{i,t-1}\left(\widehat{h}\right)$ and, therefore, $m_{i,t}\left(\widehat{h}\right) = m_{i,t-1}\left(\widehat{h}\right)$ implying that (iii) holds at date t for every $i \in \mathcal{I}^H$. Now consider $i \in \mathcal{I}^L$ and recall that $m_{i,t}\left(\widehat{h}\right) = \widetilde{m}_t\left(w_0^L | \alpha_i, P\right)\left[l_{i,t-1}\left(\widehat{h}\right), m_{i,t-1}\left(\widehat{h}\right), a_{i,t}\left(\widehat{h}\right)\right]$. Since $p_t > p_r$ and $\omega^L\left(p_{t-1}, \alpha_i\right) \geq \omega^H\left(p_{t-1}, \alpha_i\right)$, then $G\left[\omega^L\left(p_{t-1}, \alpha_i\right), \alpha_i\right] \cdot m_{i,t-1}\left(\widehat{h}\right) < r \cdot a_{i,t-1}\left(\widehat{h}\right) = a_{i,t}\left(\widehat{h}\right)$. Therefore, $m_{i,t}\left(\widehat{h}\right) = G\left[\omega^L\left(p_{t-1}, \alpha_i\right), \alpha_i\right] \cdot m_{i,t-1}\left(\widehat{h}\right)$ and then (iii) holds at t for every $i \in \mathcal{I}^L$. ■

Lemma A.3 Let $x_i^t \in H_i^t$ and $h' = h^*\left(x_i^t, \widehat{\gamma}\right)$. If $\chi_{j,t}\left(h'\right) = 0$ for all $j \in \mathcal{I}$, then $R_{i,t}\left(h'\right) = R^*\left[p_t, w_{i,t}\left(h'\right), \alpha_i\right]$.

Proof of Lemma A.3: Since $v_{j,t}\left(h'\right) = v^*$ for every $j \in \mathcal{S}_i$, $v_{i,t}^e\left(h'\right) = v^*$ and $\chi_{i,t}\left(h'\right) = 0$, then workers trained by firm $j \in \mathcal{S}_i$ do not move to firm i. Thus, $\widehat{e}_{i,t}\left(h'\right) = 0$. Since $v_{i,t}\left(h'\right) = v^*$, $v_{j,t}^e\left(h'\right) = v^*$ and $\chi_{j,t}\left(h'\right) = 0$ for every $j \in \mathcal{D}_i$, workers trained by firm i do not move either after the training process has ended. Thus, $\widehat{s}_{i,t}\left(h'\right) = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,t}\left(h'\right)$. Then, $R_{i,t}\left(h'\right) = R^*\left(p_t, w_{i,t}\left(h'\right), \alpha_i\right) \cdot \frac{m_{i,t}(h')}{a_{i,t}(h')} + r \cdot \frac{b_{i,t}(h')}{a_{i,t}(h')}$. Since $w_{i,0}\left(h'\right) \in [\underline{w}, \overline{w})$, then $R^*\left(p_t, w_{i,t}\left(h'\right), \alpha_i\right) \geq r$ if $t \leq T-1$ and $R^*\left(p_t, w_{i,t}\left(h'\right), \alpha_i\right) = r$ if $t \geq T$. Suppose $i \in \mathcal{I}^H$. Since $m_{i,t}\left(h'\right) = a_{i,t}\left(h'\right)$ if $t \leq T-1$, then $R_{i,t}\left(h'\right) = R^*\left[p_t, w_{i,t}\left(h'\right), \alpha_i\right]$. Suppose $i \in \mathcal{I}^L$. Since $m_{i,0}\left(h'\right) = a_{i,0}\left(h'\right)$ and $R^*\left[p_t, w_{i,t}\left(h'\right), \alpha_i\right] = r$ for every $t \geq 1$, then $R_{i,t}\left(h'\right) = R^*\left[p_t, w_{i,t}\left(h'\right), \alpha_i\right]$. ■

Let $e_{i,t}\left(\overline{x^t}\right) = e_{i,t}\left(h\right)$ where $\overline{h_i^t}\left(h\right) = \overline{x_i^t}$ and $m_{i,t}\left(\overline{x^t}\right)$, $v_{i,t}\left(\overline{x^t}\right)$, $v_{i,t}^e\left(\overline{x^t}\right)$ and $\chi_{i,t}\left(\overline{x^t}\right)$ be similarly defined. If $w \leq w_{i,t}\left(\widehat{h}\right)$, set $\Omega_{i,t}^1\left(w\right) = I_{i,t}^1\left(w\right)$. If $w > w_{i,t}\left(\widehat{h}\right)$, define

$$\Omega_{i,t}^1\left(w\right) = \left\{ \overline{x^t} \in I_{i,t}^1\left(w\right) : \begin{array}{c} \frac{e_{i,t}\left(\overline{x^t}\right)}{e_{i,t}\left(\widehat{h}\right)} \geq 1, m_{i,t}\left(\overline{x^t}\right) = m_{i,t}\left(\widehat{h}\right), b_{i,t}\left(\overline{x^t}\right) = b_{i,t}\left(\widehat{h}\right) \\ \overline{v}_{j,t}\left(\overline{x^t}\right) = v_{j,t}^e\left(\overline{x^t}\right) = v^* \,\&\, \chi_{j,t}\left(\overline{x^t}\right) = 0,\ \forall j \neq i \end{array} \right\}$$

Let $\widetilde{m}_t\left(T, \delta\right)\left(x_i^t\right) = \widetilde{m}_t\left(T, \delta\right)\left(l_{i,t-1}\left(h\right), m_{i,t-1}\left(h\right), a_{i,t}\left(h\right)\right)$, $\widetilde{e}_t\left(\alpha_i\right)\left(x_i^t\right) \equiv \widetilde{e}_t\left(\alpha_i\right)\left(l_{i,t-1}\left(h\right), m_{i,t-1}\left(h\right), a_{i,t}\left(h\right)\right)$ for $h_i^t\left(h\right) = x_i^t$ and

$$\Omega_{i,t}\left(x_i^t\right) = \left\{ x^t \in I_{i,t}\left(x_i^t\right) : \begin{array}{l} \sum_{j \in \mathcal{S}_i \cap \mathcal{I}^H} \frac{(1-\alpha_j) \cdot \widetilde{m}_t(T,\delta)\left(x_j^t\right)}{v^*+c} > \widetilde{e}_t\left(\alpha_i\right)\left(x_i^t\right) \\ \overline{v}_{j,t}\left(x^t\right) = v_{j,t}^e\left(x^t\right) = v^*, \frac{(1-\alpha_j) \cdot \widetilde{m}_t(T,\delta)\left(x_j^t\right)}{v^*+c} < \lambda \cdot \widehat{l}_{j,t-1}\left(\overline{x^{t-1}}\right)\ \forall j \neq i, \end{array} \right\}$$

Proof of Proposition 4.2: The proof consists in three steps. The first step shows that the payoff of playing $f_i \in \left\{f_i^H, f_i^L\right\}$ at information set $I_{i,t}\left(x_i^t\right)$ is

$$E_{\widehat{\Psi}_{I_{i,t}\left(x_i^t\right)}}\left[\pi_{i,t}\left(f_i, \widehat{\gamma}_{-i}\right| \alpha_i, P)\right] = a_{i,t}\left(h'\right) \cdot \sum_{k=t}^{\infty} \beta^{k+1-t} \prod_{\tau=t}^{k} R^*\left(p_\tau, w_{i,\tau}\left(h'\right), \alpha_i\right) \tag{23}$$

where $h' = h^*\left(\widehat{x}^t, \widehat{\gamma}\right)$. The second step shows no other strategy can yield a larger payoff to firm i. The third step argues that no young worker has a profitable deviation either at date t.

<u>First step</u>: Suppose $x_i^t = h_i^t\left(\widehat{h}\right)$. Then, $h' = \widehat{h}$ and $E_{\widehat{\Psi}_{I_{i,t}\left(x_i^t\right)}}\left[\pi_{i,t}\left(f_i, \widehat{\gamma}_{-i}\right| \alpha_i, P)\right] = \sum_{k=t}^{\infty} \beta^{k+1-t} a_{i,k}\left(\widehat{h}\right)$. By Proposition 4.1, $a_{i,k+1}\left(\widehat{h}\right) = R^*\left(p_k, w_{i,k}\left(\widehat{h}\right), \alpha_i\right) \cdot a_{i,k}\left(\widehat{h}\right)$ for every $k \geq t$ and we conclude that (23) holds.

Suppose $x_i^t \neq h_i^t\left(\widehat{h}\right)$. To show that (23) holds, it suffices to argue that for every $k \geq t$

$$R_{i,k}\left(h'\right) = R^*\left(p_k, w_{i,k}\left(h'\right), \alpha_i\right) \tag{24}$$

First, I show that $\chi_{j,k}\left(h'\right) = 0$ for all $k \geq t+1$ and $j \in \mathcal{I}$. That (24) holds for any $k \geq t+1$ then follows by Lemma A.3. Since $w_{j,k}\left(h'\right)$ is the cutoff value of the young workers' strategy at date $k \geq t$ for every $j \in \mathcal{I}$, then $\widehat{l}_{j,k}\left(h'\right) = \widetilde{l}_k\left(\alpha_j\right)\left[l_{j,k-1}\left(h'\right), m_{j,k-1}\left(h'\right), a_{j,k}\left(h'\right)\right] = \frac{\alpha_j}{w_{j,k}\left(\widehat{h}\right)} \cdot m_{j,k}\left(h'\right)$ for every $k \geq t$ and $j \in \mathcal{I}$. Therefore, $\chi_{j,k}\left(h'\right) = 0$ for every $k \geq t+1$ if and only if $\frac{1-\alpha_j}{\alpha_j} \cdot \frac{w_{j,k}(h')}{v^*+c} \cdot \frac{m_{j,k}(h')}{m_{j,k-1}(h')} \leq \lambda$. If AW holds the latter always holds by Lemmas A.2 and 3.2. If $\mu^H = 1$, $\frac{c}{\theta} = \overline{w}_2$, $p_0 \leq \frac{p^*(\alpha_j)}{r} \cdot \frac{\alpha_j}{1-\alpha_j} \cdot \frac{\overline{w}_2+c}{\overline{w}_1} \cdot \lambda$ and P is nonincreasing, then $\frac{1-\alpha_j}{\alpha_j} \cdot \frac{w_{j,k}(h')}{v^*+c} \cdot \frac{m_{j,k}(h')}{m_{j,k-1}(h')} \leq \frac{1-\alpha_j}{\alpha_j} \cdot \frac{w_{j,k}(h')}{v^*+c} \cdot p_0 \cdot q\left(\overline{w}_1, \alpha_j\right) \leq \frac{1-\alpha_j}{\alpha_j} \cdot \frac{w_{j,k}(h')}{v^*+c} \cdot p_0 \cdot \frac{r}{p^*(\alpha_j)} < \lambda$.

Finally, I show (24) holds at date t. Suppose x_i^t is such that $\chi_{i,t}\left(h'\right) = 0$. Since $\chi_{j,t}\left(h'\right) = 0$ on $\Omega_{i,t}\left(x_i^t\right)$, Lemma A.3, once again, implies that $R_{i,t}\left(h'\right) = R^*\left[p_t, w_{i,t}\left(h'\right), \alpha_i\right]$. Now, suppose x_i^t is such that $\chi_{i,t}\left(h'\right) \neq 0$. Consider first a firm $j \in \mathcal{S}_i$. Since $v_{\widetilde{i},t}^e\left(h'\right) = v^*$, $\chi_{\widetilde{i},t}\left(h'\right) = 0$ for every $\widetilde{i} \in \mathcal{D}_j$ such that $\widetilde{i} \neq i$ and $v_{j,t}\left(h'\right) = v^*$ on $\Omega_{i,t}\left(x_i^t\right)$ and $v_{i,t}^e\left(h'\right) = v^*$, then $s_{j,t}\left(h'\right) = \frac{(1-\alpha_j)\cdot\widetilde{m}_t(T,\delta)\left(x_j^t\right)}{v^*+c}$ and workers trained by firm j who receive an offer from firm i move to firm i. Since $\sum_{j \in \mathcal{S}_i \cap \mathcal{I}^H} \frac{(1-\alpha_j)\cdot\widetilde{m}_t(T,\delta)\left(x_j^t\right)}{v^*+c} > \widetilde{e}_t\left(\alpha_i\right)\left(l_{i,t-1}\left(h'\right), m_{i,t-1}\left(h'\right), a_{i,t}\left(h'\right)\right)$ on $\Omega_{i,t}\left(x_i^t\right)$, then $\widehat{e}_{i,t}\left(h'\right) = e_{i,t}\left(h'\right) = \frac{1-\alpha_i}{v^*} \cdot m_{i,t}\left(h'\right) - (1+\theta)\cdot\lambda\cdot l_{i,t-1}\left(h'\right)$. Consider now $j \in \mathcal{D}_i$. Since $\chi_{j,t}\left(h'\right) = 0$ and $v_{j,t}^e\left(h'\right) = v^*$ for every $j \in \mathcal{D}_i$ on $\Omega_{i,t}\left(x_i^t\right)$, then $\widehat{s}_{i,t}\left(h'\right) = s_{i,t}\left(h'\right) = \lambda \cdot l_{i,t-1}\left(h'\right)$. Therefore, $\widehat{s}_{i,t}\left(h'\right) + \frac{\widehat{e}_{i,t}(h')}{1+\theta} = \frac{1-\alpha_i}{v^*+c} \cdot m_{i,t}\left(h'\right)$ and $R_{i,t}\left(h'\right) = p_t \cdot \left(\frac{\alpha_i}{w_{i,t}\left(\widehat{h}\right)}\right)^{\alpha_i} \left(\frac{1-\alpha_i}{v^*+c}\right)^{1-\alpha_i} \cdot \frac{m_{i,t}\left(\widehat{h}\right)}{a_{i,t}\left(\widehat{h}\right)} + r \cdot \frac{b_{i,t}\left(\widehat{h}\right)}{a_{i,t}\left(\widehat{h}\right)} = R^*\left(p_t, w_{i,t}\left(h'\right), \alpha_i\right)$ by an argument analogous to the one used in Lemma A.3.

Then, (24) holds for any $k \geq t$ and we conclude (23) also holds for any $x_i^t \neq h_i^t\left(\widehat{h}\right)$.

<u>Second step</u>: Suppose firm i has a profitable deviation at information set $I_{i,t}\left(x_i^t\right)$. Then there exists a strategy $f \in \mathbb{F}_i$ such that $E_{\widehat{\Psi}_{I_{i,t}\left(x_i^t\right)}}\left[\pi_{i,t}\left(f, \widehat{\gamma}_{-i}\right| \alpha_i, P)\right] > E_{\widehat{\Psi}_{I_{i,t}\left(x_i^t\right)}}\left[\pi_{i,t}\left(f_i, \widehat{\gamma}_{-i}\right| \alpha_i, P)\right]$. Since

$$E_{\widehat{\Psi}_{I_{i,t}\left(x_i^t\right)}}\left[\pi_{i,t}\left(f, \widehat{\gamma}_{-i}\right| \alpha_i, P)\right] = a_{i,t}\left(h'\right) \cdot \sum_{k=t}^{\infty} \beta^{k+1-t} \prod_{\tau=t}^{k} R_{i,\tau}\left(h''\right)$$

where $h'' = h^*\left(\widehat{x}^t, f, \widehat{\gamma}_{-i}\right)$, there is $k \geq t$ such that $R_{i,k}\left(h''\right) > R^*\left(p_k, w_{i,k}\left(h'\right), \alpha_i\right)$. Since $R^*\left(p_k, w_{i,k}\left(h'\right), \alpha_i\right) \geq r$, it follows that $R_{i,k}\left(h''\right) > r$ and $m_{i,k}\left(h''\right) > 0$. This implies that $\widehat{l}_{i,k}\left(h''\right) > 0$ and $\widehat{s}_{i,k}\left(h''\right) + \frac{\widehat{e}_{i,k}(h'')}{1+\theta} > 0$. Then $w_{i,k}\left(h''\right) \geq w_{i,k}\left(h'\right)$, $\widehat{s}_{i,k}\left(h''\right) \cdot \left(v_{i,k}\left(h''\right) - v^*\right) \geq 0$, $\widehat{e}_{i,k}\left(h''\right) \cdot \left(v_{i,k}^e\left(h''\right) - v^*\right) \geq 0$ and $w_{i,k}\left(h''\right) \cdot \widehat{l}_{i,k}\left(h''\right) + \left(v_{i,k}\left(h''\right) + c\right) \cdot \widehat{s}_{i,k}\left(h''\right) + v_{i,k}^e\left(h''\right) \cdot \widehat{e}_{i,k}\left(h''\right) \leq m_{i,k}\left(h''\right)$. Therefore,

$$R_{i,k}\left(h''\right) \leq \max_{l,s,e\geq 0} p_k \cdot l^{\alpha_i}\left(s+\frac{e}{1+\theta}\right)^{1-\alpha_i} \quad s.t. \begin{cases} w_{i,k}\left(h'\right)\cdot l+\left(v^*+c\right)\cdot s+v^*\cdot e=1 \\ s\cdot\left(v_{i,k}\left(h''\right)-v^*\right)\geq 0 \text{ and } e\cdot\left(v^e_{i,k}\left(h''\right)-v^*\right)\geq 0 \end{cases}$$

But this implies that $R_{i,k}\left(h''\right) \leq R^*\left(p_k, w_{i,k}\left(h'\right), \alpha_i\right)$, a contradiction.

Third step: Suppose a young worker has a profitable deviation at information set $I^1_{i,t}\left(w\right)$. Suppose $w = w_{i,t}\left(\widehat{h}\right)$. If he rejects, he obtains $\overline{w}_1+\beta\cdot\overline{w}_2$. Since $w_{i,t}\left(\widehat{h}\right)+\beta\cdot\left(\frac{s_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}\cdot\left[v^*-\overline{w}_2\right]+\overline{w}_2\right)=\overline{w}_1+\beta\cdot\overline{w}_2$, a contradiction is reached. Suppose $w \neq w_{i,t}\left(\widehat{h}\right)$. Let $h^w = h^*\left(\widehat{x^t},\widehat{\gamma}\right)$. Suppose $w > w_{i,t}\left(\widehat{h}\right)$. He obtains $w+\beta\cdot\left(\frac{s_{i,t+1}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)}\cdot\left[v^*-\overline{w}_2\right]+\overline{w}_2\right)$ if he accepts and $\overline{w}_1+\beta\cdot\overline{w}_2$ if he rejects. If one shows that $\frac{s_{i,t+1}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)} \geq \frac{s_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}$, no profitable deviation can exist because $w+\beta\cdot\left(\frac{s_{i,t+1}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)}\cdot\left[v^*-\overline{w}_2\right]+\overline{w}_2\right) > w_{i,t}\left(\widehat{h}\right)+\beta\cdot\left(\frac{s_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}\cdot\left[v^*-\overline{w}_2\right]+\overline{w}_2\right)$. Since $s_{i,t+1}\left(h^w\right) = \frac{1-\alpha_i}{v^*+c}\cdot\frac{m_{i,t+1}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)}$ and $s_{i,t+1}\left(\widehat{h}\right) = \frac{1-\alpha_i}{v^*+c}\cdot\frac{m_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}$, it suffices to show that $\frac{m_{i,t+1}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)} \geq \frac{m_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}$. Since every other young worker accepts, $\widehat{l}_{i,t}\left(h^w\right) = l_{i,t}\left(h^w\right)$. Since $v_{j,t}\left(\widehat{x^t}\right) = v^e_{j,t}\left(\widehat{x^t}\right) = v^*$ and $\chi_{j,t}\left(\widehat{x^t}\right) = 0$ for every $j \in \mathcal{I}$, then $\widehat{e}_{j,t}\left(h^w\right) = 0$ for every $j \in \mathcal{I}$. Hence, $\widehat{s}_{i,t}\left(h^w\right) = s_{i,t}\left(h^w\right) = \frac{v^*\cdot e_{i,t}\left(h^w\right)}{v^*+c} > \frac{v^*\cdot e_{i,t}\left(\widehat{h}\right)}{v^*+c} = \widehat{s}_{i,t}\left(\widehat{h}\right)$ and $\widehat{l}_{i,t}\left(h^w\right) = \frac{m_{i,t}\left(h^w\right)-v^*\cdot e_{i,t}\left(h^w\right)}{w} < \frac{m_{i,t}\left(\widehat{h}\right)-v^*\cdot e_{i,t}\left(\widehat{h}\right)}{w_{i,t}\left(\widehat{h}\right)} = \widehat{l}_{i,t}\left(\widehat{h}\right)$. If $m_{i,t+1}\left(h^w\right) = \delta_{t+1}\cdot a_{i,t+1}\left(h^w\right)$, where $\delta_{t+1}=1$ if $t \neq T-1$ and $\delta_T = \delta$, $\frac{m_{i,t+1}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)} = \delta_{t+1}\cdot\left[p_t\cdot\left(\frac{\widehat{s}_{i,t}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)}\right)^{1-\alpha_i}+r\cdot\frac{b_{i,t}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)}\right] \geq \delta_{t+1}\cdot\left[p_t\cdot\left(\frac{\widehat{s}_{i,t}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}\right)^{1-\alpha_i}+r\cdot\frac{b_{i,t}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}\right] = \frac{\delta_{t+1}\cdot a_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)} \geq \frac{m_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}$. If $t \geq T$ and $m_{i,t+1}\left(h^w\right) < a_{i,t+1}\left(h^w\right)$, then $m_{i,t+1}\left(h^w\right) = m_{i,t}\left(h^w\right) = m_{i,t}\left(\widehat{h}\right) = m_{i,t+1}\left(\widehat{h}\right)$ and again $\frac{m_{i,t+1}\left(h^w\right)}{\widehat{l}_{i,t}\left(h^w\right)} \geq \frac{m_{i,t+1}\left(\widehat{h}\right)}{\widehat{l}_{i,t}\left(\widehat{h}\right)}$. One concludes there is no profitable deviation when $w > w_{i,t}\left(\widehat{h}\right)$. When $w < w_{i,t}\left(\widehat{h}\right)$, every other young worker rejects and so $\widehat{l}_{i,t}\left(h^w\right)=0$ and $s_{i,t+1}\left(h^w\right)=0$ because $q_{i,t}\left(h^w\right)=0$. If he accepts, he obtains $w+\beta\cdot\overline{w}_2 < \overline{w}_1+\beta\cdot\overline{w}_2$ and so no profitable deviation exists with $w < w_{i,t}\left(\widehat{h}\right)$ either. ■

C. Proofs of Section 6

Proof of Lemma 6.1: Suppose AD holds and $m \leq \frac{p_s\cdot D(p_s)}{r}$. Consider the function $H:\Re_+\times\left[0,\frac{p_s\cdot D(p_s)}{r}\right]\to\Re$ defined by $H\left(p,m\right)=q\left(\omega^H\left(p\right)\right)\cdot m-D\left(p\right)$. By Lemma 3.3 and assumption AD it follows that H is continuous and strictly increasing in both m and p. By AD, $\lim_{p\to\infty}H\left(p,m\right)>0$. Since H is continuous, to show that $H\left(p,m\right)=0$ has a solution it suffices to show that there exists p such that $H\left(p,m\right)<0$. Notice that $\omega^H\left(\frac{p_s}{r}\right)$ is the unique solution to $w=\overline{w}_1-\beta\cdot\frac{1-\alpha}{\alpha}\cdot\frac{w}{v^*+c}\cdot\frac{p_s}{r}\cdot q\left(w\right)\cdot\left(v^*-\overline{w}_2\right)$. Since $\omega\left(1\right)$ is also a solution to that equation, it follows that $\omega^H\left(\frac{p_s}{r}\right)=\omega\left(1\right)$. Hence $H\left(\frac{p_s}{r},m\right)=q\left(\omega^H\left(\frac{p_s}{r}\right)\right)\cdot m-D\left(\frac{p_s}{r}\right)=q\left(\omega\left(1\right)\right)\cdot m-D\left(\frac{p_s}{r}\right)<\frac{r}{p_s}\cdot m-D\left(p_s\right)\leq 0$ where the last inequality follows from the assumption that $m\leq\frac{p_s\cdot D(p_s)}{r}$. By the intermediate value theorem there exists $p>\frac{p_s}{r}$ such that $H\left(p,m\right)=0$. Since H is strictly increasing in its first argument, the solution is unique. Therefore there exists a function $\mathbb{P}$:$\left[0,\frac{p_s\cdot D(p_s)}{r}\right]\to\left(\frac{p_s}{r},\infty\right)$ such that $H\left[\mathbb{P}\left(m\right),m\right]=0$. Notice that $H\left[p_r,m\right]<H\left[p_r,\frac{p_r\cdot D(p_r)}{r}\right]=0=H\left[\mathbb{P}\left(m\right),m\right]$ if and only if $m<\frac{p_r\cdot D(p_r)}{r}$. Hence $\mathbb{P}\left(m\right)>p_r$ if and only if $m<\frac{p_r\cdot D(p_r)}{r}$, as desired. ■

Lemma A.4 *Let $\frac{c}{\theta}\geq\overline{w}_1$. Suppose $0<m<\frac{p_s\cdot D(p_s)}{r}$ and $P\left(m\right)\cdot D\left(\mathbb{P}\left(m\right)\right)\geq\frac{p_s\cdot D(p_s)}{r}$. Equation (14) has a unique solution $p_{T-1}=D^{-1}\left(q\left[\omega\left(\frac{p_s\cdot D(p_s)}{r\cdot m}\right)\right]\cdot m\right)\geq Max\left\{p_s,\mathbb{P}\left(m\right)\right\}$ and $p_{T-1}\cdot D\left(p_{T-1}\right)\geq\frac{p_s\cdot D(p_s)}{r}$.*

Proof of Lemma A.4: Let $\frac{c}{\theta} > \overline{w}_1$. Suppose $0 < m < \frac{p_s \cdot D(p_s)}{r}$ and $\mathbb{P}(m) \cdot D(\mathbb{P}(m)) \geq \frac{p_s \cdot D(p_s)}{r}$. Since $0 < \frac{p_s \cdot D(p_s)}{r} \leq \mathbb{P}(m) \cdot D(\mathbb{P}(m))$, then $w_{T-1} = \omega\left(\frac{p_s \cdot D(p_s)}{r \cdot m}\right) \geq \omega\left(\frac{\mathbb{P}(m) \cdot D(\mathbb{P}(m))}{m}\right) = \omega^H(\mathbb{P}(m))$. Hence, $\omega^H(\mathbb{P}(m)) \leq w_{T-1} \leq \omega(1)$. Since $q(w_{T-1}) \cdot m \leq q\left(\omega^H(\mathbb{P}(m))\right) \cdot m = D(\mathbb{P}(m)) \leq D\left(\frac{p_s}{r}\right)$ then $p_{T-1} = D^{-1}(q(w_{T-1}) \cdot m)$ is well defined because D has an inverse on $\left[0, \frac{p_s}{r}\right]$. Clearly, p_{T-1} solves (14) and $p_{T-1} \geq \mathbb{P}(m)$. Uniqueness follows from AD. Since $q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot m}\right)\right] \cdot m$ is strictly increasing in m,

$$D(p_{T-1}) = q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot m}\right)\right] \cdot m < q[\omega(1)] \cdot \frac{p_s \cdot D(p_s)}{r} = D(p_s)$$

for all $m < \frac{p_s \cdot D(p_s)}{r}$. Therefore, $p_{T-1} > p_s$ and $p_{T-1} \geq \max\{p_s, \mathbb{P}(m)\}$.

I show $p_{T-1} \cdot D(p_{T-1}) \geq \frac{p_s \cdot D(p_s)}{r}$ by reduction to the absurd. Suppose $p_{T-1} \cdot D(p_{T-1}) < \frac{p_s \cdot D(p_s)}{r}$. Then,

$$\begin{aligned} B\left(\omega^H(p_{T-1}), p_{T-1}\right) &= 0 = \overline{w}_1 - \frac{1-\alpha}{\alpha} \cdot \frac{w_{T-1}}{v^* + c} \cdot \frac{p_s \cdot D(p_s)}{r \cdot m} \cdot \beta \cdot (v^* - \overline{w}_2) - \omega_{T-1} \\ &< \overline{w}_1 - \frac{1-\alpha}{\alpha} \cdot \frac{w_{T-1}}{v^* + c} \cdot p_{T-1} \cdot q(w_{T-1}) \cdot \beta \cdot (v^* - \overline{w}_2) - \omega_{T-1} = B(w_{T-1}, p_{T-1}) \end{aligned}$$

and since $B(\cdot, p_{T-1})$ is decreasing in its first argument, it follows that $w_{T-1} < \omega^H(p_{T-1})$. Hence, $\omega^H(\mathbb{P}(m)) < \omega^H(p_{T-1})$ which implies that $\mathbb{P}(m) > p_{T-1}$, a contradiction since $D(p_{T-1}) = q(w_{T-1}) \cdot m \leq D(\mathbb{P}(m))$ and D is decreasing in p. It follows that $p_{T-1} \cdot D(p_{T-1}) \geq \frac{p_s \cdot D(p_s)}{r}$, as desired. ■

Proof of Proposition 6.1: First I consider the existence of an IE and then I turn to its uniqueness.

EXISTENCE: There are two cases to consider depending on the value of $\mathbb{P}(a_0) \cdot D(\mathbb{P}(a_0))$.

Case (i): $\mathbb{P}(a_0) \cdot D(\mathbb{P}(a_0)) \geq \frac{p_s \cdot D(p_s)}{r}$

By Lemma A.4, the equation $q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r a_0}\right)\right] \cdot a_0 = D(p)$ has a unique solution p_0 and $\frac{p_s \cdot D(p_s)}{r} \leq p_0 \cdot D(p_0)$. Let $\widehat{\delta} = \frac{p_s \cdot D(p_s)}{r \cdot p_0 \cdot D(p_0)}$. Clearly, $\widehat{\delta} \in [0, 1]$. Let $T_s = 1$ and define $P = \{p_0, p_s, p_s, ...\}$. Clearly, $P \in \Sigma$ and is non-increasing because $p_0 > p_s \geq p_r$. In addition, $p_0 \cdot q\left[\omega^H\left(\widehat{\delta} \cdot p_0\right)\right] > p_s \cdot q(\omega(1)) = r$. By Proposition 4.2, $\left(\widehat{\gamma}, \widehat{\Psi}\right)$ is a *PBESO. By construction, the output market clears at date zero. At any other date $t \geq 1$, $q[w_{i,t}(h^*)] \cdot m_{i,t}(h^*) = q(\omega(1)) \cdot m_1^H(h^*) = q(\omega(1)) \cdot \frac{p_s \cdot D(p_s)}{r} = D(p_s) = D(p_t)$. Then, there is an IE.

Case (ii): $\mathbb{P}(a_0) \cdot D(\mathbb{P}(a_0)) < \frac{p_s \cdot D(p_s)}{r}$

Let $y_t = a_0$ and

$$y_{t+1} = \begin{cases} \mathbb{P}(y_t) \cdot D(\mathbb{P}(y_t)) & \text{if } \mathbb{P}(y_t) \cdot D(\mathbb{P}(y_t)) < \frac{p_s \cdot D(p_s)}{r} \text{ and } y_t < Min\left\{\frac{p_r \cdot D(p_r)}{r}, \frac{p_s \cdot D(p_s)}{r}\right\} \\ \frac{p_s \cdot D(p_s)}{r} & \text{otherwise} \end{cases}$$

Let τ be the first date t such that $y_t \geq \frac{p_s \cdot D(p_s)}{r}$. Clearly, $\tau \geq 1$ because $a_0 < \frac{p_s \cdot D(p_s)}{r}$. I show that τ is finite. Suppose not. Then $y_t = \mathbb{P}(y_{t-1}) \cdot D(\mathbb{P}(y_{t-1})) = \mathbb{P}(y_{t-1}) \cdot q\left(\omega^H(\mathbb{P}(y_{t-1}))\right) \cdot y_{t-1} \geq r \cdot y_{t-1}$ because $y_{t-1} \leq \frac{p_r \cdot D(p_r)}{r}$ implies that $\mathbb{P}(y_{t-1}) \geq p_r$. It follows that $y_t \geq r^t \cdot a_0$ which implies that $y_t \to \infty$, a contradiction. Thus, τ is finite. Let $T_s = \tau$ and P be the sequence with $p_t = \mathbb{P}(y_t)$ for all $t < T_s$, $p_t = p_s$ for all $t \geq T_s$ and with $p_{T_s - 1}$ as the solution to $q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s - 1}}\right)\right] \cdot y_{T_s - 1} = D(p)$. Since $y_t < \frac{p_r \cdot D(p_r)}{r}$ for all $t < T_s$, then $p_t > p_r$ for all $t < T_s$. To prove that $P \in \Sigma$, I shall show that $p_{T_s - 1}$ is well defined and $p_{T_s - 1} > p_r$.

By definition of T_s, $y_{T_s - 1} < \frac{p_s \cdot D(p_s)}{r}$. Suppose $y_{T_s - 1} < \frac{p_r \cdot D(p_r)}{r}$. Then $\mathbb{P}(y_{T_s - 1}) \cdot D(\mathbb{P}(y_{T_s - 1})) \geq \frac{p_s \cdot D(p_s)}{r}$ and by Lemma A.4 $p_{T_s - 1}$ is well defined, $p_{T_s - 1} > p_s \geq p_r$ and $\widehat{\delta} = \frac{p_s \cdot D(p_s)}{r \cdot p_{T_s - 1} \cdot D(p_{T_s - 1})} \in [0, 1]$. Therefore,

$p_{T_s-1} \cdot q\left[\omega^H\left(\widehat{\delta} \cdot p_{T_s-1}\right)\right] = p_{T_s-1} \cdot q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s-1}}\right)\right] > p_s \cdot q\left(\omega\left(1\right)\right) = r$. Suppose $y_{T_s-1} \geq \frac{p_r \cdot D(p_r)}{r}$. Since $q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s-1}}\right)\right] \cdot y_{T_s-1}$ is strictly increasing in y_{T_s-1}, it follows that $q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s-1}}\right)\right] \cdot y_{T_s-1} < q\left[\omega\left(1\right)\right] \cdot \frac{p_s \cdot D(p_s)}{r} = D\left(p_s\right)$ for all $y_{T_s-1} < \frac{p_s \cdot D(p_s)}{r}$. Hence, by AD there is a unique p_{T_s-1} such that $q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s-1}}\right)\right] \cdot y_{T_s-1} = D\left(p_{T_s-1}\right)$. In addition, $p_{T_s-1} > p_s \geq p_r$. Therefore, $p_{T_s-1} \cdot q\left[\omega^H\left(\widehat{\delta} \cdot p_{T_s-1}\right)\right] = p_{T_s-1} \cdot q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s-1}}\right)\right] > p_s \cdot q\left(\omega\left(1\right)\right) = r$ and $p_{T_s-1} \cdot D\left(p_{T_s-1}\right) = p_{T_s-1} \cdot q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s-1}}\right)\right] \cdot y_{T_s-1} > p_r \cdot D\left(p_r\right) \geq \frac{p_s \cdot D(p_s)}{r}$, where the last inequality holds because $p_r \geq \frac{p_s}{r}$ and $D\left(p_r\right) > D\left(p_s\right)$. Hence, $\widehat{\delta} = \frac{p_s \cdot D(p_s)}{r \cdot p_{T_s-1} \cdot D(p_{T_s-1})} \in [0,1]$.

It follows that $P \in \Sigma$ and is non-increasing and $p_{T_s-1} \cdot q\left[\omega^H\left(\widehat{\delta} \cdot p_{T_s-1}\right)\right] > r$. By Proposition 4.2, for $\widehat{\delta}$ and $T = T_s - 1$, $\left(\widehat{\gamma}, \widehat{\Psi}\right)$ is a *PBESO. Finally, I shall show that the output market clears at every $t \geq 0$. Since $m_0^H\left(h^*\right) = a_0 = y_0$ and $m_t^H\left(h^*\right) = p_{t-1} \cdot q\left(\omega^H\left(p_{t-1}\right)\right) \cdot a_{t-1} = \mathbb{P}\left(a_{t-1}\right) \cdot D\left(\mathbb{P}\left(a_{t-1}\right)\right)$, then $m_t^H\left(h^*\right) = y_t$ for all $0 \leq t \leq T_s - 1$. Hence,

$$q_t^H\left(h^*\right) = q\left[\omega^H\left(p_t\right)\right] \cdot m_t^H\left(h^*\right) = q\left[\omega^H\left(p_t\right)\right] \cdot y_t = D\left(p_t\right) \qquad \forall 0 \leq t < T_s - 1$$

and the output market clears. At date $T_s - 1$,

$$\begin{aligned} q_{T_s-1}^H\left(h^*\right) &= q\left[\omega^H\left(\widehat{\delta} \cdot p_{T_s-1}\right)\right] \cdot m_{T_s-1}^H\left(h^*\right) \\ &= q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot m_{T_s-1}^H(h^*)}\right)\right] \cdot m_{T_s-1}^H\left(h^*\right) = q\left[\omega\left(\frac{p_s \cdot D(p_s)}{r \cdot y_{T_s-1}}\right)\right] \cdot y_{T_s-1} = D\left(p_{T_s-1}\right) \end{aligned}$$

Finally, at any date $t \geq T_s$, $q_t^H\left(h^*\right) = q\left[\omega\left(1\right)\right] \cdot m_{T_s}^H\left(h^*\right) = q\left[\omega\left(1\right)\right] \cdot \frac{p_s \cdot D(p_s)}{r} = D\left(p_s\right)$ as desired.

UNIQUENESS: Suppose there is another IE $\left\{\widetilde{P}, 1, \left(\widetilde{\gamma}, \widetilde{\Psi}\right)\right\}$ for some $\widetilde{\delta} \in [0,1]$ and $\widetilde{T} \geq \widetilde{T}_s$. Let $\widetilde{h} = h^*\left(\widetilde{\gamma}\right)$.

Suppose $\widetilde{T} > T$. For every $i \in \mathcal{I}$, $\left(\widetilde{f}_{i,t}^H, \widetilde{\sigma}_{i,t}^H\right) = \left(\widehat{f}_{i,t}^H, \widehat{\sigma}_{i,t}^H\right)$ for all $t \leq T - 2$ and $m_{T-1}^H\left(\widetilde{h}\right) = m_{T-1}^H\left(h^*\right) < \frac{p_s D(p_s)}{r}$. It follows that $\widetilde{p}_t = p_t$ for all $t \leq T - 2$. Since $\widetilde{p}_t = p_s = p_t$ for all $t \geq \widetilde{T}$ then $\widetilde{P} \neq P$ if and only if there exists $T - 1 \leq t \leq \widetilde{T} - 1$ such that $\widetilde{p}_t \neq p_s$ and $\widetilde{p}_t \geq p_r$. From the construction of the price equilibrium sequence P, it follows that either $\mathbb{P}\left(m_{T-1}^H\left(h^*\right)\right) \cdot D\left(\mathbb{P}\left(m_{T-1}^H\left(h^*\right)\right)\right) \geq \frac{p_s D(p_s)}{r}$ or $m_{T-1}^H\left(\widetilde{h}\right) \geq \frac{p_r D(p_r)}{r}$. Suppose $m_{T-1}^H\left(\widetilde{h}\right) \geq \frac{p_r D(p_r)}{r}$. Since $\widetilde{p}_{T-1} \geq p_r$, then $m_{T-1}^H\left(\widetilde{h}\right) \leq \frac{p_r D(p_r)}{r}$. Hence, $m_{T-1}^H\left(\widetilde{h}\right) = \frac{p_r D(p_r)}{r}$ and $\mathbb{P}\left(m_{T-1}^H\left(\widetilde{h}\right)\right) = p_r$. Since $p_r \cdot D\left(p_r\right) \geq \frac{p_s D(p_s)}{r}$ then it is always the case that $\mathbb{P}\left(m_{T-1}^H\left(\widetilde{h}\right)\right) \cdot D\left(\mathbb{P}\left(m_{T-1}^H\left(\widetilde{h}\right)\right)\right) \geq \frac{p_s D(p_s)}{r}$. Therefore, $m_{T-1}^H\left(\widetilde{h}\right) \geq \frac{p_s D(p_s)}{r}$ for all $t \geq T$. If $m_{\widetilde{T}-1}^H\left(\widetilde{h}\right) > \frac{p_s D(p_s)}{r}$, then

$$D\left(\widetilde{p}_{\widetilde{T}-1}\right) = q\left(w_{i,\widetilde{T}-1}\left(\widetilde{h}\right)\right) \cdot m_{\widetilde{T}-1}^H\left(\widetilde{h}\right) = q\left(\frac{p_s D(p_s)}{r \cdot m_{\widetilde{T}-1}^H\left(\widetilde{h}\right)}\right) \cdot m_{\widetilde{T}-1}^H\left(\widetilde{h}\right) > q\left[\omega\left(1\right)\right] \cdot \frac{p_s D(p_s)}{r} = D\left(p_s\right)$$

Then $\widetilde{p}_{\widetilde{T}-1} < p_s$ and $\widetilde{p}_{\widetilde{T}-1} \cdot q\left(w_{\widetilde{T}-1}\left(\widetilde{h}\right)\right) < p_s \cdot q\left[\omega\left(1\right)\right] = r$. Therefore, $\left(\widetilde{\gamma}, \widetilde{\Psi}\right)$ is not a *PBESO of $\Gamma\left(\widetilde{P}\right)$, a contradiction. If $m_{\widetilde{T}-1}^H\left(\widetilde{h}\right) = \frac{p_s D(p_s)}{r}$, instead, then $m_t^H\left(\widetilde{h}\right) = \frac{p_s D(p_s)}{r}$ for all $t \geq T$. Therefore, $w_{i,t}\left(\widetilde{h}\right) = \omega\left(1\right)$ for every $i \in \mathcal{I}$ which implies that $\widetilde{p}_t = p_s = p_t$ for all $t \geq T$. Hence $\widetilde{P} = P$.

Suppose $\widetilde{T} < T$. Clearly, $\left(\widetilde{f}_{i,t}^H, \widetilde{\sigma}_{i,t}^H\right) = \left(\widehat{f}_{i,t}^H, \widehat{\sigma}_{i,t}^H\right)$ for all $t \leq \widetilde{T} - 2$, and $i \in \mathcal{I}$ and $m_{\widetilde{T}-1}^H\left(\widetilde{h}\right) = m_{\widetilde{T}-1}^H\left(h^*\right) < \frac{p_s D(p_s)}{r}$. Since $\widetilde{\delta} = \frac{p_s D(p_s)}{r \cdot \widetilde{p}_{\widetilde{T}-1} \cdot D\left(\widetilde{p}_{\widetilde{T}-1}\right)}$ and $\widetilde{\delta} \leq 1$, then

$$\widetilde{p}_{\widetilde{T}-1} \cdot q\left(w_{i,\widetilde{T}-1}\left(\widetilde{h}\right)\right) = \widetilde{p}_{\widetilde{T}-1} \cdot q\left(\omega^H\left(\widetilde{\delta} \cdot \widetilde{p}_{\widetilde{T}-1}\right)\right) = \widetilde{p}_{\widetilde{T}-1} \cdot D\left(\widetilde{p}_{\widetilde{T}-1}\right) \geq \frac{p_s D(p_s)}{r}$$

It follows that

$$\mathbb{P}\left(m_{\widetilde{T}-1}^H\left(\widetilde{h}\right)\right) \cdot D\left(\mathbb{P}\left(m_{\widetilde{T}-1}^H\left(\widetilde{h}\right)\right)\right) \geq \widetilde{p}_{\widetilde{T}-1} \cdot q\left(\omega^H\left(\widetilde{\delta} \cdot \widetilde{p}_{\widetilde{T}-1}\right)\right) \geq \frac{p_s D(p_s)}{r} > \mathbb{P}\left(m_{\widetilde{T}-1}^H\left(h^*\right)\right) \cdot D\left(\mathbb{P}\left(m_{\widetilde{T}-1}^H\left(h^*\right)\right)\right)$$

a contradiction since $m^{H}_{\widetilde{T}-1}\left(\widetilde{h}\right) = m^{H}_{\widetilde{T}-1}\left(h^{*}\right)$. Thus, $\widetilde{P} = P \Rightarrow \left(\widetilde{f}^{H}_{i,t}, \widetilde{\sigma}^{H}_{i,t}\right) = \left(\widehat{f}^{H}_{i,t}, \widehat{\sigma}^{H}_{i,t}\right) \forall t \geq 0$. ■

Proof of Proposition 6.2: By Proposition 3.7, it follows that $i \in \mathcal{I}^{L}$ and for every $j \in \mathcal{I}^{H}$, $w_{i,t}\left(h^{*}\right) = w^{L}\left(p_{t}\right) < \omega^{H}\left(p_{t}\right) = w_{j,t}\left(h^{*}\right)$ for all $1 \leq t \leq T-1$. ■

References

[1] A. Alchian, Uncertainty, Evolution, and Economic Theory, *J. of Political Economy*, **58** (1950), 211 - 221.

[2] W. B. Arthur, Competing Technologies, Increasing Returns and Lock-In By Historical Events, *The Economic Journal* **99** (1989), 116-131.

[3] F. Allen and D. Gale, "Comparing Financial Systems", The MIT Press, Cambridge, Massachusets, 2000.

[4] P. Beker, Are Inefficient Entrepreneurs Driven Out of the Market?, *Journal of Economic Theory* **114** (2004), 329-344.

[5] P. Beker, Retained Earnings Dynamic, Internal Promotions and Walrasian Equilibrium, Working Paper AD 2004-14, Ivie, also available at http://www.ivie.es/downloads/docs/04/wpad-14.pdf.

[6] D. Bernhardt and D. Scoones, Promotion, Turnover, and Preemptive Wage Offers, *The American Economic Review*, 83 (1993), 771-791.

[7] L. Blume and D. Easley, Optimality and Natural Selection in Markets, *Journal of Economic Theory **107** (2002), 95-135.*

[8] P. Doeringer and M. Piore, "Internal Labor Markets and Manpower Analysis", D.C. Heath and Company, Massachusetts, 1971.

[9] M. Feldman and C. Gilles, An Expository Note on Individual Risk without Aggregate Uncertainty, *Journal of Economic Theory* **35** (1985), 26-32.

[10] M. Friedman, "Three Essays on the State of Economic Science", University of Chicago Press, Chicago,1953.

[11] J. Malcomson, Work Incentives, Hierarchy, and Internal Labor Markets, *J. Polit. Economy* **92** (1984), 486-507.

[12] R. Nelson and S. Winter, "An evolutionary Theory of Economic Change", The Belknap Press of Harvard University Press, Cambridge, 1982.

[13] J. Stiglitz and A. Weiss, Credit Rationing in Markets with Imperfect Information, *The American Economic Review* (1981), 393 - 410.