Work Practices, Incentives for Skills, and Training*

Suraj Prasad[†] and Hien Tran

June 15, 2012

Abstract

We study a worker's incentives to acquire skills that cannot be verified by a third party. We do this within the context of recent innovative work practices, where jobs have become more flexible. Because flexible jobs are difficult to verify by a court, standard approaches in the literature, where firms use contracts on jobs to commit to reward workers for skills, may not be feasible. We suggest an alternative approach to induce skills: firms promote workers to challenging jobs (where the returns to skills are high) only when they acquire *both* firm specific skills and general skills. This promotion scheme reveals information about the general skill to competing firms which in turn allows the firm to commit to reward firm specific skills. So, general training in our framework plays the same role as contracts on jobs in other papers in the literature; it helps a firm to commit to reward firm specific skills. Consistent with empirical evidence on innovative work practices, we find that firms are more likely to train workers in general skills when jobs are difficult to contract on (i.e. when jobs are flexible and job classifications are broad).

1 Introduction

The past few decades have seen significant changes in the organization of work in firms.[1] While the details of these changes differ across firms and industries there are some common features. First, there is a greater emphasis on *continuous skill development and training* at all levels of an organization. Some of these skills are firm specific such as interpersonal skills with team members and customers. Other

*We would like to thank Jed DeVaro, Denise Doiron, Richard Holden, Clare Leaver, James Malcomson, Kieron Meagher, Hodaka Morita, Bill Schworm, Margaret Stevens, Joel Watson and participants at the Economic Theory Workshop, UNSW, Australian National University, California State University, East Bay, University of New South Wales, and University of Oxford.

[†]School of Economics, University of New South Wales. `s.prasad@unsw.edu.au`

[1]Examples of surveys where these trends are documented are Osterman (1994), Osterman (2000), and Lawler, Mohrman, and Ledford (1995). Also see Cappelli and Neumark (2001), Black and Lynch (2004) for evidence on the adoption of innovative work practices. O'Toole and Lawler (2006) provide a detailed description of recent work practices and Lindbeck and Snower (2000) and Ichniowski and Shaw (2003) survey some of the empirical literature in this area.

skills are general such as solving problems in the production process and making decisions. Second, jobs have become more *flexible* in response to changes in technology, markets and customers needs. Firms rely less on stable job descriptions and more on flexible work-assignment descriptions. And these dynamic descriptions change as assignments are completed. Third, workers are offered *challenging jobs* that utilize their skills as they progress through their careers. The objective of our paper is to understand how firms design incentives for workers to acquire skills in light of these recent trends in the workplace.

The literature on incentives for skills suggests a few approaches. The first approach is to write a contract that rewards workers for skills. The problem with this approach is that skills such as developing a relationship with a customer or making improvements to the production process are often difficult to verify by a third party and thus difficult to contract on. A second approach is for firms to write an *indirect contract* on a variable related to skills, say jobs (Kahn and Huberman (1988) and Prendergast (1993)). This helps a firm to commit ex-ante to rewarding workers in some jobs, which it can use ex-post by assigning workers to jobs based on skills. But when jobs change over time in response to changes in technology, markets and customers needs, writing these contracts may not be feasible. In particular, it may be difficult to describe a job ex-ante and for courts to verify these descriptions ex-post. Firms can also commit to a tournament where a fixed fraction of workers from a pool are promoted and rewarded (Carmichael (1983), Malcomson (1984)). But tournaments can ruin cooperation between members of a team. Finally, firms can rely on their reputations to commit to reward workers (MacLeod and Malcomson (1988), Baker, Gibbons, and Murphy (1994), Levin (2003)). This approach, however, requires firms to be sufficiently patient and requires workers to coordinate their punishments.

We make two contributions in this paper. First, we suggest an alternative mechanism to induce skills, where workers are promoted to challenging jobs (where the returns to skills are high) only when they acquire both firm specific skills (such as interpersonal skills with team members and customers) and general skills (such as problem solving skills and decision making). Unlike the existing approaches above, our mechanism does not rely on ex-ante contracts on skills or on jobs, and can be used in firms where jobs are flexible. What our mechanism does need, instead, is that competing firms in the same industry (which have access to the same technology and are familiar with product markets) *can* verify job assignments (as in Waldman (1984)). Second, building on our mechanism, we suggest a new rationale for why firms pay for general training; general skills can be used to commit to reward workers for firm specific skills. Consistent with empirical evidence on innovative work practices, our theory suggests that firms are more likely to provide training when jobs are flexible and when job classifications are broad. These implications for training which rely on the inability to contract on jobs are distinct from standard explanations in the literature which are based on other labor market imperfections.

To understand how our mechanism works, it is useful to start with a paper by Prendergast (1993). He considers the role of promotions in inducing firm specific skills. His framework has two main features. First, though firms cannot write a contract on skills, they *can* write a contract on jobs, *before* skills have been acquired. Second, jobs vary in their returns to skills and a promotion is an assignment to a job where the returns to skills are higher. Firms can then split the promise to reward a worker for skills into two parts. The first promise is to promote the worker for skills (promotion promise) and the second promise is to pay a higher wage for a promotion that covers the cost of skills (wage promise). Because contracts are written on jobs, firms can always set the reward for a promotion high enough and keep their wage promise. And by setting the reward for a promotion low enough, relative to the output gain from promoting a skilled worker, a firm can keep its promotion promise. So, contracts written on jobs help a firm to commit to reward firm specific skills.

Our mechanism builds on the framework by Prendergast (1993). Like in Prendergast (1993), a promise to reward workers for skills is split into two parts in our framework: a promise to promote the worker for skills and a promise to pay a higher wage for a promotion that covers the cost of skills. But, in contrast to Prendergast (1993), we assume that it is not feasible, for firms to write contracts ex-ante on jobs. Thus the firm cannot directly set rewards for a promotion. Instead, rewards for a promotion are determined by the information about skills that a promotion scheme conveys to competing firms in a labor market. Promoting the worker based on a general skill (such as problem solving and decision making) that is sufficiently valuable ensures that the reward for a promotion covers the cost of skills. So the firm can keep its wage promise. Promoting the worker based on a firm specific skill (such as learning to work with team members and customers) that is sufficiently valuable, on the other hand, helps the firm keep its promotion promise. This is because the firm specific skill, as in Waldman (1984), increases the output gains from promotion without getting reflected in wage offers by the market. To keep both promises, a worker is promoted only when he successfully acquires both skills. Thus general skills that cannot be verified by a third party play exactly the same role as contracts on jobs in the Prendergast framework: they help a firm to commit to reward firm specific skills.

We fix this commitment role of general skills in inducing firm specific skills and derive some additional results. First, promotions to challenging jobs depend on both the firm specific and the general skill. Second, this promotion scheme induces workers to invest in both firm specific and general skills, even when the skills are not complements in the production function. This result stands in contrast to other papers which emphasize the role of complementarities in multiskilling (Lindbeck and Snower (2000), Gibbs and Levenson (2002)).[2] Third, this promotion scheme

[2]Starting with Adam Smith (Smith (1937)) there is a large literature which emphasizes the benefits of specialization (Roy (1950) and Rosen (1978)). Other explanations for multiskilling are

sometimes leads to over-investment in the firm specific skill.

The fact that general skills can be used to commit to reward firm specific skills also suggests a new explanation for why firms provide general training to workers. To make this point, we modify our model and allow firms to incur a fixed cost and offer a (publicly observable) training program to workers upfront. This program in turn allows workers to make investments in the general skill, which only the current employer observes. A promotion that is contingent on both types of skills selectively reveals information about the general skill to a labor market only when the worker acquires the firm specific skill. This helps a firm commit to a reward, which along with a worker's access to skills induces a worker to invest in firm specific skills. Because general training plays the same role as ex-ante contracts on jobs in rewarding firm specific skills, firms are more likely to provide training when jobs are difficult to contract on. This result is consistent with evidence on innovative work practices (Ichniowski, Shaw, and Prennushi (1997), Ichniowski and Shaw (2003)) where firms are more likely to train workers when jobs are flexible and job classifications are broad. This implication for training, which relies on the inability to contract on jobs, is distinct from other explanations that rely on other labor market imperfections (Acemoglu and Pischke (1999) for example).

Two key features of our model, promotions and asymmetric information between current employers and the labor market also play a role in Waldman (1984).[3] In his framework, firms face the following tradeoff when deciding on a promotion. Promotions increase output but they also signal ability to a labor market which increases wages. As a result, there is always some under-promotion relative to the efficient level. Waldman (1984) also shows that firm specific skills, which are exogenous in his framework, increase output from a promotion without altering wages, and thus make promotions more likely.

A few papers that build on Waldman's work also consider investment in skills. Scoones and Bernhardt (1998) focus on skills that are *verifiable by a labor market* whereas our focus is on non-verifiable skills.[4] Zabojnik and Bernhardt (2001) study how firm characteristics such as size and profitability influence investments in skills and wages. But they do not study investments in firm specific skills which is an important part of our paper. Ghosh and Waldman (2010) study a worker's incentive to exert effort (as opposed to investing in skills) under standard promotion schemes and up or out schemes. Finally, DeVaro, Ghosh, and Zoghi (2008) study promotion

based on the extent of the market (Smith (1937)) and coordination costs (Becker and Murphy (1992)).

[3]A number of other papers assume asymmetric information about a workers productivity. Examples are Waldman (1990), Gibbons and Katz (1991), Milgrom and Oster (1987), Bernhardt (1995), Owan (2004) and Chang and Wang (1995). Also see DeVaro and Waldman (2011) for empirical evidence on the signalling role of promotions and Schonberg (2007) for empirical evidence on asymmetric employer learning.

[4]Promotions do not convey information about skills in their framework though they do convey information about exogenous ability. So promotions as in Waldman (1984) only depend on the firm specific skill. Firms, also do not have an incentive to pay for general training.

patterns within the context of labor market discrimination where general and firm specific skills are acquired in fixed proportions.

Our paper is also related to work on multitasking and multiskilling (Holmstrom and Milgrom (1991), Baker (1992)). This literature suggests that the provision of incentives in multitask settings is difficult because incentives have to be balanced across tasks which differ in the preciseness of their performance measures. Our model, in contrast, suggests that jobs with different types of skills help in the provision of incentives. In this sense, our paper is related to Carmichael and MacLeod (1993) where multiskilling allows firms to commit to retain workers after workers suggest a labor saving innovation. Finally, our result where workers are rewarded if and only if they are successful at both skills has a similar flavor to results in MacDonald and Marx (2001).

2 Model

We build on a model by Prendergast (1993), with a firm where the worker is currently working and an outside labor market with at least two competing firms. All of the agents are risk neutral. In the analysis that follows, we often refer to the current employer as *the firm* and competing firms as *the labor market*. The model has three parts. The first part deals with the skill acquisition process. The second part relates skills with output. The final part specifies the timing of the game and the information that the players have.

Consider the acquisition of skills first. There are two skills that the worker can acquire, a firm specific skill and a general skill, with each skill having two outcomes. The outcome for the firm specific skill is denoted by i with $i \in \{0, f\}$ where 0 denotes failure and f denotes success at the firm specific skill. Similarly, the outcome for the general skill is denoted by j with $j \in \{0, g\}$ where 0 denotes failure and g denotes success for the general skill. Thus, there are four possible outcomes, failure at both skills $(0, 0)$, success only at the firm specific skill $(f, 0)$, success only at the general skill $(0, g)$ and success at both skills (f, g). The worker can exert effort that is privately observable to influence the probability of success on each skill.[5] To simplify notation, we assume that the worker chooses these probabilities directly. The probability of acquiring the firm specific skill is $p_f \in [\underline{p}, \bar{p}]$ and the probability of acquiring the general skill is $p_g \in [\underline{p}, \bar{p}]$ and these probabilities are independent of one another. To ensure that beliefs of the labor market can be determined by Baye's rule, we assume that $\underline{p} > 0$ and $\bar{p} < 1$. In the analysis that follows we refer to p_f as the level of investment in the firm specific skill and p_g as the level of investment in

[5]This implies that acquiring skills is a risky activity. We think it is reasonable within the context of our examples in the introduction such as developing relationships with clients and acquiring technical skills.

the general skill. We assume that the worker has a cost function given by $C(p_f, p_g)$ defined over the domain $[0,1] \times [0,1]$ that is strictly positive, strictly increasing, convex and differentiable in the interior of the domain. We also assume that there exists a lower bound $b > 0$ such that for all p_f and p_g in the interior of the domain, $min\{\frac{\partial C(p_f,p_g)}{\partial p_f}, \frac{\partial C(p_f,p_g)}{\partial p_g}\} \geq b$. The following definition is also useful in stating the main results of the paper.

Definition 1. *A worker specializes in a skill if* $max\{p_f, p_g\} > \underline{p}$ *and* $min\{p_f, p_g\} = \underline{p}$.

We say that a worker multiskills if $max\{p_f, p_g\} > \underline{p}$ and he does not specialize in a skill.

Next, consider how skills are related to output. The relationship between skills and output depends on the job assignment. There are two jobs, an easy or a low skilled job, job L, where the returns to skills are low and a challenging or a high skilled job, job H, where the returns to skills are high. Output for the four possible outcomes is as follows. When the outcome is $(0,0)$, output is 0 for both the firm and the labor market. When the worker is only successful at the firm specific skill, output for the firm is λy_L for job L and λy_H for job H with $0 < y_L < y_H$ and where $\lambda \in (0,1)$ measures the relative value of firm specific skills to general skills, whereas the labor market gets an output of 0.[6] When the worker is only successful at the general skill, output for both the firm and the labor market is $(1-\lambda)y_L$ for job L and $(1-\lambda)y_H$ for job H. Finally for the outcome (f,g), output for the firm is given by $y_L + \Delta$ for job L and $y_H + \Delta$ for job H with $\Delta \geq 0$, whereas output for the labor market is given by $(1-\lambda)y_J$ where $J \in \{L, H\}$. The notation Δ measures the extent to which investments in both the skills are complementary. In other words, when we fix a job, Δ is the cross partial of expected output with respect to p_f and p_g.[7] The table below summarizes the relationship between skills and output for the firm.

Outcomes	(0,0)	(f,0)	(0,g)	(f,g)
Output from Job L	0	λy_L	$(1-\lambda)y_L$	$y_L + \Delta$
Output from Job H	0	λy_H	$(1-\lambda)y_H$	$y_H + \Delta$

Finally, the timing in the model is as follows. The worker chooses the probabilities of success for both skills which induces a probability distribution over outcomes.

[6] Restricting λ to lie in the open rather than the closed interval simplifies the analysis. All of our results go through if we consider the end points of the interval.

[7] We could allow for some of the gains in complementarities to be captured by a competing firm in the labor market. If we did make this assumption, then complementarities in investments would make it easier for a firm to keep its wage promise in equilibrium.

The outcomes for both of these skills are then realized. The firm observes these outcomes and decides on the job assignment of the worker. Competing firms in the labor market then observe the job assignment, update their beliefs and offer a wage that is contingent on the job assignment. We focus on symmetric wage offers by competing firms. The firm then decides whether to match the labor market's offer for a job. If the firm matches an offer made by the labor market, we assume that the worker stays with the firm. We also assume, that the firm incurs a cost $F > 0$ if it loses the worker.[8] This assumption ensures that it is optimal for firms to match offers even when the offers equal the value of output at a given outcome.

In this setting a strategy for a worker is given by the vector (p_f, p_g). The firm's strategy is denoted by $(\sigma(x), m)$ where $x = (x_{0,0}, x_{f,0}, x_{0,g}, x_{f,g})$ with $x_{i,j} \in \{L, H\}$, where σ is a probability distribution over x, and where m is a function which specifies whether the firm matches offers for a given outcome, job assignment and wage offer for the job. The labor market observes job assignments and for each job assigns beliefs to each of the four outcomes. The beliefs associated with each of the outcomes for job L are given by the vector of probabilities $\mu^L = (\mu^L_{0,0}, \mu^L_{f,0}, \mu^L_{0,g}, \mu^L_{f,g})$ and the beliefs associated with each outcome for job H are given by the vector of probabilities $\mu^H = (\mu^H_{0,0}, \mu^H_{f,0}, \mu^H_{0,g}, \mu^H_{f,g})$. Wages offered by the labor market are denoted by w_L when job L is observed and w_H when job H is observed.

3 Efficiency

We start our analysis by characterizing efficient levels of investment for both skills. This characterization serves as a benchmark. Let (p_f^{eff}, p_g^{eff}) denote the efficient level of investment. Then (p_f^{eff}, p_g^{eff}) is the optimal solution to the following problem.

$$\underset{p_f \in [\underline{p}, \bar{p}], p_g \in [\underline{p}, \bar{p}]}{Max} \quad p_f(1-p_g)\lambda y_H + (1-p_f)p_g(1-\lambda)y_H + p_f p_g(y_H + \Delta) - C(p_f, p_g)$$

To make the comparison easier with the section on equilibrium skill acquisition, we restrict our attention in this section to a linear cost function, $C(p_f, p_g) = c_f p_f + c_g p_g$ where c_f and c_g are strictly positive constants. We also make the following assumption to ensure that the efficient solution always sets at least one of the skill levels above the minimum level.

Assumption 1. $y_H > c_f + c_g$

The following proposition characterizes the efficient skill acquisition levels. The proofs of all of the propositions and lemmas are in the Appendix.

[8] There are various ways to interpret the cost F. It could be a firing cost, the cost of hiring a new worker, or firm specific capital that is lost.

Proposition 1. *Let $C(p_f, p_g) = c_f p_f + c_g p_g$ where c_f and c_g are strictly positive constants. Then the efficient levels of skill acquisition are given by*

$$p_f^{eff} = \begin{cases} \underline{p} \ \ \textit{if} \ \ \Delta \leq \dfrac{c_f}{\bar{p}} \ \textit{and} \ \ \lambda \in (0, \dfrac{c_f - \bar{p}\Delta}{y_H}) \\ [\underline{p}, \bar{p}] \ \textit{if} \ \ \Delta \leq \dfrac{c_f}{\bar{p}} \ \textit{and} \ \ \lambda = \dfrac{c_f - \bar{p}\Delta}{y_H} \\ \bar{p} \ \ \textit{if} \ \ \Delta \leq \dfrac{c_f}{\bar{p}} \ \textit{and} \ \ \lambda \in (\dfrac{c_f - \bar{p}\Delta}{y_H}, 1) \\ \bar{p} \ \ \textit{if} \ \ \Delta > \dfrac{c_f}{\bar{p}} \end{cases}$$

and

$$p_g^{eff} = \begin{cases} \bar{p} \ \ \textit{if} \ \ \Delta > \dfrac{c_g}{\bar{p}} \\ \bar{p} \ \ \textit{if} \ \ \Delta \leq \dfrac{c_g}{\bar{p}} \ \textit{and} \ \ \lambda \in (0, 1 - \dfrac{c_g - \bar{p}\Delta}{y_H}) \\ [\underline{p}, \bar{p}] \ \textit{if} \ \ \Delta \leq \dfrac{c_g}{\bar{p}} \ \textit{and} \ \ \lambda = 1 - \dfrac{c_g - \bar{p}\Delta}{y_H} \\ \underline{p} \ \ \textit{if} \ \ \Delta \leq \dfrac{c_g}{\bar{p}} \ \textit{and} \ \ \lambda \in (1 - \dfrac{c_g - \bar{p}\Delta}{y_H}, 1) \end{cases}$$

Proposition 1 highlights two factors that determine efficient investment levels: complementarities between skills and the relative importance of both skills. When complementarities are high, investment levels are at their maximum for both skills. When complementarities are low, the relative value of each skill plays an important role. If both skills are important (for intermediate values of λ), efficiency requires a worker to choose the highest investment level for both skills. If one skill is more important relative to another (λ close to 0 or 1), efficiency requires an extreme form of specialization, the highest investment level for the more valuable skill and the lowest investment level for the less valuable skill. Notice that these corner solutions arise because of the linear specification for the cost function. Figure 1 depicts the efficient level of investment with λ on the horizontal axis and p_f^{eff} and p_g^{eff} on the vertical axis when complementarities are not too high.

4 Equilibrium

The objective in this section is to compare equilibrium investment levels for skills with the efficient counterparts above. The equilibrium concept we use is a Perfect

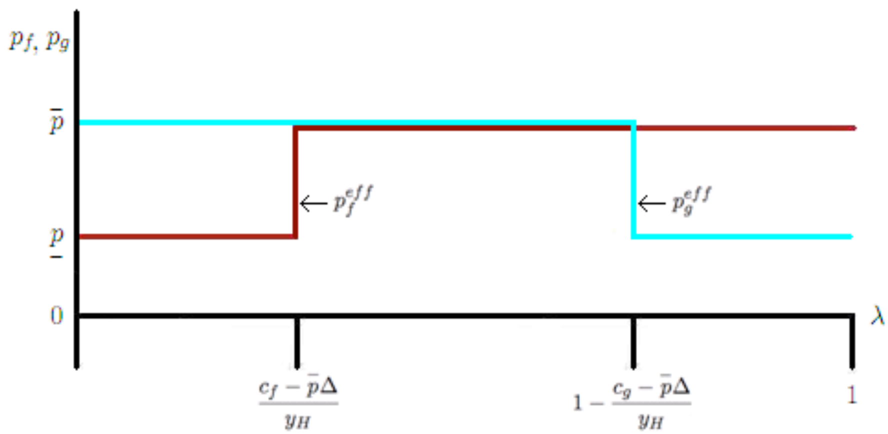

Figure 1: Efficient Investment Levels.

Bayesian Equilibrium and focus on equilibria where investments for at least one skill are strictly above the minimum level.[9] To start solving the equilibrium, let us start at the last stage where the firm decides whether it wants to match offers or not. Because the firm is better informed than competing firms about skill outcomes and because of the presence of firm specific skills, a competing firm in the labor market can only successfully raid workers with a wage that exceeds the worker's value to it. Thus the best that an outside firm can do is to earn an expected profit of zero and there are several wage offers by competing firms which satisfy this condition, each providing different incentives for the worker to acquire skills. To pin down wage offers, we follow Ghosh and Waldman (2010) and allow for the possibility of a firm trembling and not matching offers when it should. In particular, given a job assignment and given an offer by an outside firm we assume that the firm trembles at the outcome where it has the lowest output.[10] A firm in the labor market then makes an offer that equals the output that it gets from job H for that particular outcome.[11] As an example, suppose the firm's promotion strategy is (L, L, H, H).

[9]However, when computing the equilibrium we do check for deviations where a worker chooses the minimum investment level for both skills.

[10]This is similar to the notion of a proper equilibrium (Myerson (1978)). An alternative way to pin down offers is to assume that firm's cost of losing the worker is sufficiently large. This way, when a competing firm bids its expected value conditional on the job it observes, the offer is always matched. The results using this alternative approach are qualitatively similar when we consider pure strategies only. For the case of mixed strategies, it is possible that the worker specializes in the general skill.

[11]If two outcomes yield equal output for the firm and are both consistent with a job assignment (when $\lambda = \frac{1}{2}$), the labor market picks the outcome where its own output is lower and offers the value of its output as the wage.

Then w_L is the labor market's output in job H associated with the outcome $(0,0)$ which is 0, and w_H is the labor market's output in job H corresponding to the outcome $(0,g)$ which is $(1-\lambda)y_H$. The following lemma says that rewards are always associated with promotions in a skill acquisition equilibrium.

Lemma 1. *In a skill acquisition equilibrium, $w_H \geq w_L$.*

The reasoning for this Lemma is the following. To induce skills there must be some outcome for which a worker is assigned to the lower skilled job (a pooling strategy where a worker is always assigned to the higher skilled job cannot induce skills). But if w_L was strictly greater than w_H, he firm could always promote the worker, get at least as much output, pay a lower wage and do strictly better.

Proposition 2. *In a skill acquisition equilibrium, the firm assigns a worker to H with positive probability if and only if the outcome (f,g) is realized. Furthermore, if $\underline{p}y_H < b$, the worker never specializes in a skill.*

Proposition 2 says that in a skill acquisition equilibrium, a worker is promoted if and only if he successfully acquires both skills. As a consequence, if $\underline{p}$ is sufficiently small then both types of skills are always acquired together in a skill acquisition equilibrium. The interesting feature of this result is that it holds even if the skills are not complements ($\Delta = 0$). A simple way to understand Proposition 2 is to restrict attention to pure strategies for the firm and eliminate strategies that cannot be part of a skill acquisition equilibrium. Pooling strategies cannot be part of a skill acquisition equilibrium. Also, promotion strategies that are more likely to promote a worker for outcomes with lower output gains cannot be part of an equilibrium. This leaves us with four possible strategies to check: (L,H,L,H), (L,L,H,H), (L,H,H,H) and (L,L,L,H). Consider (L,H,L,H) which induces specialization in the firm specific skill. This strategy cannot be part of a skill acquisition equilibrium because the reward for a promotion equals 0 and the firm cannot keep its wage promise. Next, consider (L,L,H,H) which induces specialization in the general skill. This strategy cannot be part of a skill acquisition equilibrium because the reward for a promotion, $(1-\lambda)y_H$, is too high relative to the output gain $(1-\lambda)(y_H - y_L)$, and the firm cannot keep its promotion promise at $(0,g)$. Similarly (L,H,H,H) can be ruled out using the arguments above. Thus the only possible pure strategy equilibrium with skills is (L,L,L,H). When $\underline{p}y_H < b$, the worker does not specialize in either skill because the marginal benefit from specializing in a skill, which is bounded above by $\underline{p}y_H$, is always less than the marginal cost.

Since the worker must invest in both skills in a skill acquisition equilibrium, there is a possibility for over investment of skills relative to the efficient level. To make this inefficiency as stark as possible, we restrict our attention to a cost function which is linear in the worker's investment. Linear costs, guarantee us corner solutions in equilibrium where skills are acquired at either the maximum or minimum possible level. Also define $\bar{c} = max\{\frac{c_f + c_g}{\underline{p} + \bar{p}}, \frac{c_f}{\bar{p}}, \frac{c_g}{\bar{p}}\}$. The following proposition characterizes investments in both skills in equilibrium when costs are linear.

Proposition 3. *Let $C(p_f, p_g) = c_f p_f + c_g p_g$ where c_f and c_g are strictly positive constants. Then there exists a skill acquisition equilibrium in pure strategies with investment levels $(\bar{p}, \bar{p})$ and a promotion scheme (L, L, L, H) if and only if the following conditions hold*

$$\bar{c} \leq y_H - y_L \tag{1}$$

$$\lambda \in [1 - \frac{(y_H - y_L)}{y_H}, min\{1 - \frac{\bar{c}}{y_H}, \frac{y_H}{(y_H - y_L) + y_H}\}]. \tag{2}$$

Furthermore, when $\underline{p} y_H < b$, this skill acquisition equilibrium is unique in pure strategies.

The best way to understand Proposition 3 is to compare it with Prendergast (1993). Recall that in his paper there has to be a gap between the output gains from promotion and the cost of acquiring skills, which is exactly what the inequality in (1) states. The firm then, as in Prendergast's framework, can set the reward for a promotion somewhere in between these two values to keep its promotion promise and to induce skills. The key difference in our framework is that without a court the firm loses its flexibility in setting the reward for a promotion. Instead, as condition (2) shows, the reward is determined by information about skills that a promotion conveys to the labor market. If only the firm specific skill is valuable, the reward offered by the labor market is too small relative to the cost of acquiring both skills and if only the general skill is valuable, the firm will not have an incentive to keep its promotion promise. The main point to take away from Proposition 3 is that general skills, that cannot be verified by a third party, play a similar role to contracts on jobs in Prendergast's framework. That is they help a firm to commit to reward firm specific skills.

We can also compare equilibrium promotions and skill levels with their efficient counterparts in Section 3. As in Waldman (1984), there is some under-promotion. But the pattern is different. Workers are under promoted in our setting when they do not acquire both skills together, whereas in Waldman (1984) (and in Scoones and Bernhardt (1998)) the under-promotion depends only on the firm specific skills. Next, consider investment levels. Because the worker bears all the costs of skill acquisition and the firm gets all of the benefits we would typically expect under-investment in skills relative to the efficient level in equilibrium. For λ sufficiently close to 0 or 1, this is exactly the case. What is unusual about Proposition 3 is that there can be an over-investment in skills. This can be seen by comparing the

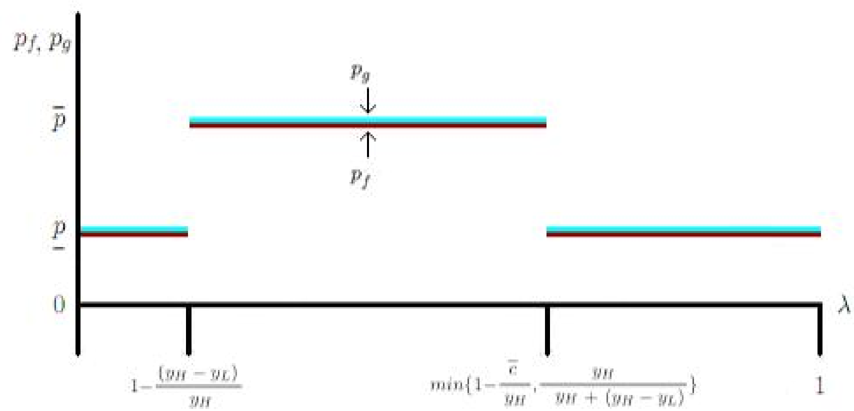

Figure 2: Equilibrium Investment Levels.

smallest λ for which a skill acquisition equilibrium exists in Figure 2 with the cutoff λ for specialization in the general skill in Figure 1. In fact when $c_f > y_L + \bar{p}\Delta$ there is over-investment in the firm specific skill.

5 Training

A key point in Proposition 3, is that general skills that are not verifiable by a third party (courts or labor markets) play exactly the same role as ex-ante contracts on jobs in Prendergast (1993): they help a firm to commit to reward specific skills. Firms can thus benefit from providing general training to workers in contrast to arguments made in Becker (1964).[12] To make this point, we modify our model and allow the firm to incur a fixed cost $C_T > 0$ and provide a general training program (that is publicly observable) before workers invest in skills. If workers are trained, then they can make investments in both the firm specific and general skills from the

[12]Becker argues that in a competitive labor market, wages get bid up to equal all of the gains in productivity from training. Thus ex-post, the firm gets none of the returns from training. The firm anticipates this ex-ante and does not incur the training cost. Becker also notes that because workers get all of the benefits from training ex-post through higher wage offers, they have an incentive ex-ante to pay for training. Taken together, his observations imply that if there is training in general skills, the worker should pay for it.

interval $[\underline{p}, \bar{p}]$. If not, they can make investments in the interval above only for the firm specific skill whereas they are forced to choose $\underline{p}$ for the general skill.[13] So the general training program gives a worker access to make investments in the general skill.[14]

In our analysis in this section, we restrict our attention to pure strategies and linear costs. We also assume that $\underline{p} y_H < b$ and that conditions (1) and (2) in Proposition 3 hold. Thus we assume that a skill acquisition equilibrium exists in pure strategies and that it is unique. Also define the lower and upper bounds of λ in condition (2) as $\underline{\lambda} = 1 - \dfrac{(y_H - y_L)}{y_H}$ and $\bar{\lambda} = min\{1 - \dfrac{\bar{c}}{y_H}, \dfrac{y_H}{(y_H - y_L) + y_H}\}$.

Now suppose the firm does not provide training. Then from Proposition 2, a skill acquisition equilibrium does not exist. The best possible equilibrium for the firm is then a pooling equilibrium where the worker is assigned to job H for every outcome. In this case, the worker's investment levels are $(\underline{p}, \underline{p})$ and the firm's expected profit is given by

$$\underline{p}(1 - \underline{p}) y_H + \underline{p}^2 (y_H + \Delta)$$

On the other hand, when the firm provides general training, there is a unique equilibrium in pure strategies given by the promotion scheme (L, L, L, H) and investment levels $(\bar{p}, \bar{p})$. In this case, the firm's expected profit is given by

$$\bar{p}(1 - \bar{p}) y_L + \bar{p}^2 (\lambda y_H + \Delta)$$

Comparing the expressions above, we notice two things. First, providing general training helps a firm to commit to reward skills, which in turn induces a worker to invest in firm specific skills. General training thus plays the same role as contracts on jobs in Prendergast (1993) and helps to induce firm specific skills. Second, notice that promotions selectively reveal information about the general skill only when a worker acquires firm specific skills. Thus, firms earn a return from general training. In fact, when $\underline{p}$ goes to 0 and $\bar{p}$ goes to 1 in the limit, a firm's return from training (net of the training cost) is $\lambda y_H + \Delta$.

Because general training and contracts on jobs are substitutes from the point of view of inducing firm specific skills, the returns from training should be higher when contracts cannot be written on jobs. To examine this point, we consider an

[13]We focus just on general training and not firm specific training because our primary interest is to point out that firms do benefit from training workers in general skills.

[14]It is worth noting that there are other ways to model access to the general skill. For example, training could reduce the cost of the general skill for the worker. This approach yields qualitatively similar results.

alternative case where firms *can* contract on jobs. In terms of the model, the only change is that the firm can credibly commit to a wage (through a contract) for a given job before the worker acquires skills. Competing firms, after observing the job assignment, can make an offer to the worker. And finally the current employer can decide to match this offer if it is strictly above the contractual wage. Given a contractual wage, we solve for an equilibrium in this alternative case.

As in the analysis without contracts, the best a competing firm can do is to make an expected profit of zero and there are multiple wage offers that satisfy this condition.[15] We pin down wage offers in the same way as in our analysis without contracts. Given a job and a promotion strategy, consider the outcome where the current employer has the smallest output. The offer by a competing firm then equals its own value of skills at that particular outcome. If this offer is less than or equal to the contracted wage, the competing firm is never successful at raiding the worker and makes an expected profit of zero. If this offer is strictly above the contracted wage, once again the competing firm makes an expected profit of zero, given that the current employer can tremble at the outcome where it has the lowest output. Also, because of the presence of firm specific skills, it must be the case in equilibrium that wages for a job are at least as large as the wage offers from competing firms.[16]

The following Proposition compares returns across the case where firms cannot contract on jobs and the case where they can.

Proposition 4. *Let $\underline{p}$ be sufficiently close to 0 and let $\bar{p}$ be sufficiently close to 1. Also let $\lambda \in (\max\{\underline{\lambda}, \frac{1}{2}\}, \bar{\lambda}]$ and $\max\{\frac{y_L}{2} + c_f, \bar{c}\} < \frac{y_H}{2}$. Then the returns from training are strictly larger when firms cannot contract on jobs.*

Proposition 4 says that when firm specific skills are sufficiently valuable, then the return from training a worker is higher when jobs cannot be contracted on. The intuition for this proposition is the following. Contracts on jobs and general training are substitutes from the point of view of inducing firm specific skills. When contracts cannot be written on jobs, the only way to extract surplus from firm specific skills is by providing general training. The assumptions that $\lambda > \frac{1}{2}$ and $\frac{y_L}{2} + c_f < \frac{y_H}{2}$ ensure that the promotion strategy (L, H, L, H) along with investment levels $(\bar{p}, \underline{p})$ is an equilibrium when jobs can be contracted on and workers are not trained. The assumption that $\frac{y_H}{2} > \bar{c}$ ensures that $\bar{\lambda} > \frac{1}{2}$ so that the interval $\lambda \in (\max\{\underline{\lambda}, \frac{1}{2}\}, \bar{\lambda}]$ is non-empty.

[15] For a job assignment, fix the outcome where the current employer has the lowest output. Then any offer which is less than or equal to the maximum of the contracted wage and the current employer's output at the given outcome is optimal.

[16] They can be strictly larger as well because firms can contractually commit to a wage for a job.

Propositions 3 and 4, taken together, yield two testable implications. First, workers who are trained in general skills will invest in firm specific skills even when jobs cannot be contracted on. Second, firms should be more likely to provide general training in settings where jobs are difficult to contract on. Both of these implications are consistent with a recent empirical literature on innovative work practices. Ichniowski, Shaw, and Prennushi (1997) and Gant, Ichniowski, and Shaw (2002) study systems of innovative work practices in steel finishing lines. Three of these innovative practices are important from the point of view of our analysis.[17] First, workers interact with one another in teams to solve production problems. Problem solving requires a worker to have general knowledge about the production process. But at the same time, this worker must also be familiar with what other workers in the establishment know. For this, workers must build connections and learn how to communicate with other co-workers. These valuable interpersonal skills are mainly firm specific. A second practice is that firms train workers in general areas like mechanical operations, the chemistry of steel, statistical process control, documenting production problems and running problem solving meetings. Third, firms use flexible jobs. That is workers are rotated across jobs and job classifications are broad. We assume that jobs are more difficult to verify by a court when they are flexible and are thus more difficult to contract on. Consistent with Proposition 3, Gant, Ichniowski, and Shaw (2002) find that in "involvement oriented firms" (with all three of the above practices) production workers communicate with a broad range of people including members of their crew, other production crews, staff and managers about a variety of issues.[18] And consistent with Proposition 4, Ichniowski, Shaw, and Prennushi (1997) find a positive and significant correlation between job rotation and training, and between broad job classifications and training.

There are other explanations in the literature for why firms pay for general training. But none of these papers consider two key elements of our analysis: the inability to contract on jobs and investments in firm specific skills. Acemoglu and Pischke (1998), Chang and Wang (1996) and Katz and Ziderman (1990) focus on informational asymmetries between the training firm and other firms in the labor market. With asymmetric information, labor markets do not bid wages up enough, giving firms some returns from general training. Malcomson, Maw, and McCormick (2003) show how apprentice contracts, where retained workers are paid a higher wage after an apprenticeship can increase the profit of the training firm. Acemoglu and Pischke (1999) show how the presence of an *exogenous and complementary* firm specific skill also gives incentives for a firm to pay for general training.[19] Firm specific skills are endogenously acquired in our framework and our results do not depend on the skills being complements. Stevens (1994) focusses on labor mar-

[17] Other practices include information sharing, careful hiring and selection, employment security, and incentive pay.

[18] In firms without these practices, workers communicate less with other workers.

[19] Complementarities between firm specific and general skills also increase the returns to general training in our framework. That is, firms are more likely to provide training when Δ is higher.

ket imperfections in a setting where skills are not entirely general or firm specific. Balmaceda (2005) and Kessler and Lulfesmann (2006) consider both general and specific investments. Their focus, however, is on the bargaining process between a firm and worker. Finally, Carmichael and MacLeod (1993) show how cross training a worker in multiple skills encourages labor saving innovations. Thus their theory also predicts a positive correlation between job flexibility and training. However, while cross training is important in steel finishing lines, it is not (as the examples above indicate) the only type of general training. Carmichael and MacLeod (1993) also do not consider firm specific skills.

6 Conclusion

In this paper, we study a worker's incentives to acquire non-verifiable skills. We do this within the context of recent work practices which make jobs difficult to verify by courts. We show that firms can induce skills by promoting workers to more challenging jobs only when they acquire both firm specific and general skills. This mechanism does not rely on court enforceable contracts. Furthermore, it suggests a new reason in the literature for why firms provide general training. General training helps a firm to commit to reward firm specific skills. Consistent with evidence on innovative work practices, we find that firms are more likely to provide training when jobs are difficult to contract on (i.e when jobs are flexible).

To emphasize the role of general skills, we abstract from worker heterogeneity in our model. We could allow for heterogeneity as in Waldman (1984) and assume that workers differ in their abilities across an interval. Adding this feature does not change our results qualitatively, provided we assume that the size of the interval is not too large relative to the cost of acquiring skills. Given that high involvement organizations spend substantial resources on carefully screening workers before hiring them, we believe that this assumption is reasonable.

Appendix

Proof of Proposition 1: The first order necessary conditions for the optimal solution are

$$\begin{pmatrix} \lambda y_H + p_g \Delta - c_f \\ (1-\lambda) y_H + p_f \Delta - c_g \end{pmatrix} = \bar{\mu}_f \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \underline{\mu}_f \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \bar{\mu}_g \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \underline{\mu}_g \begin{pmatrix} 0 \\ -1 \end{pmatrix}$$

where $\bar{\mu}_f$ is the non negative multiplier associated with the constraint $p_f \leq \bar{p}$ and $\underline{\mu}_f$ is the non negative multiplier associated with the constraint $p_f \geq \underline{p}$. Likewise, $\bar{\mu}_g$ and $\underline{\mu}_g$ are non negative multipliers associated with the inequality constraints for p_g.

In the following claim we establish that the efficient investment level is at the maximum level for at least one of the skills.

Claim The efficient level of investment has either $p_f = \bar{p}$ or $p_g = \bar{p}$ or both.

Proof Suppose the optimal solution had $p_f < \bar{p}$ and $p_g < \bar{p}$. Then the first order conditions imply $\lambda y_H + p_g \Delta - c_f \leq 0$ and $(1-\lambda) y_H + p_f \Delta - c_g \leq 0$. Adding both inequalities and rearranging we get $y_H + (p_g + p_f)\Delta \leq c_f + c_g$. But this contradicts Assumption 1. ■

Now suppose $\Delta > \frac{c_f}{\bar{p}}$ and suppose to the contrary that $p_f < \bar{p}$ at the optimum. From the claim above it follows that $p_g = \bar{p}$ at the optimum. Since $\Delta > \frac{c_f}{\bar{p}}$ the first order condition with respect to p_f can be written as $\lambda y_H + \bar{p}\Delta - c_f = \bar{\mu}_f - \underline{\mu}_f > 0$. From the complementary slackness conditions we have $p_f = \bar{p}$ at the optimum which is a contradiction. Similarly when $\Delta > \frac{c_g}{\bar{p}}$ we must have $p_g = \bar{p}$ at the optimum.

Next suppose $\Delta \leq \frac{c_f}{\bar{p}}$ and $\Delta \leq \frac{c_g}{\bar{p}}$ and consider the following partition of the interval $(0, 1)$ into five cases.

First, let $0 < \lambda < \frac{c_f - \bar{p}\Delta}{y_H}$. Rearranging the second inequality we get $\lambda y_H + \bar{p}\Delta - c_f < 0$. Combining this inequality with the constraint $p_g \leq \bar{p}$, we can write the first order condition with respect to p_f as $\lambda y_H + p_g \Delta - c_f = \bar{\mu}_f - \underline{\mu}_f < 0$. Since the multipliers are non-negative, it follows that $\underline{\mu}_f > 0$. The complementary slackness conditions imply that $p_f = \underline{p}$ at the optimum. It follows from the claim above that $p_g = \bar{p}$ at the optimum.

Second, let $\lambda = \frac{c_f - \bar{p}\Delta}{y_H}$. Rearranging, we get $-\lambda y_H = \bar{p}\Delta - c_f$. Adding y_H to both sides of the equality we get

$$(1-\lambda)y_H = y_H + \bar{p}\Delta - c_f \tag{3}$$

Using equation (3) and Assumption 1, the first order condition with respect to p_g is $(1-\lambda)y_H + p_f\Delta - c_g = \bar{\mu}_g - \underline{\mu}_g > 0$. Since the multipliers are non negative it follows that $\bar{\mu}_g > 0$. Thus $p_g = \bar{p}$ at the optimum. Likewise the first order conditions with respect to p_f yield

$$\lambda y_H + p_g\Delta - c_f = \bar{\mu}_f - \underline{\mu}_f \tag{4}$$

Substituting $\lambda y_H = c_f - \bar{p}\Delta$ and $p_g = \bar{p}$ into (4) we get

$$\lambda y_H + p_g\Delta - c_f = c_f - \bar{p}\Delta + \bar{p}\Delta - c_f = 0 \tag{5}$$

Since the objective function is linear with respect to p_f when p_g is held fixed it follows from (5) that any p_f in the interval $[\underline{p}, \bar{p}]$ is an optimal solution.

Third, let $\frac{c_f - \bar{p}\Delta}{y_H} < \lambda < 1 - \frac{c_g - \bar{p}\Delta}{y_H}$. Rearranging the first inequality we get $\lambda y_H > c_f - \bar{p}\Delta$. Substituting this inequality into the first order condition with respect to p_f we get $\lambda y_H + p_g\Delta - c_f = \bar{\mu}_f - \underline{\mu}_f > 0$. Since the multipliers are non-negative, it follows that $\bar{\mu}_f > 0$. The complementary slackness conditions imply that $p_f = \bar{p}$ at the optimum. Similar reasoning can be used to show that $p_g = \bar{p}$ at the optimum.

Fourth, let $\lambda = 1 - \frac{c_g - \bar{p}\Delta}{y_H}$. Using reasoning similar to the second case we can show that $p_f = \bar{p}$ at the optimum along with any p_g from the interval $[\underline{p}, \bar{p}]$.

Finally for the fifth case suppose $1 - \frac{c_g - \bar{p}\Delta}{y_H} < \lambda < 1$. Using reasoning similar to the first case we can show that $p_f = \bar{p}$ and $p_g = \underline{p}$ at the optimum.■

Proof of Lemma 1: Notice that (H, H, H, H) cannot be part of a skill acquisition equilibrium because the worker's wage does not depend on the outcome. Thus in a skill acquisition equilibrium, the firm must assign the worker to a low job for at least one of the outcomes. Suppose to the contrary that $w_L > w_H$, then the firm can deviate from L to H for that particular outcome, get at least the same level of output and pay a lower wage, which is a contradiction. ■

Proof of Proposition 2: The proof is divided into a series of claims.

Claim 1 Suppose the firm promotes a worker for some outcome (i, j) with positive probability and suppose the output difference across jobs is strictly higher for some other outcome (i', j'). Then in equilibrium, the uniquely optimal action for the firm at (i', j') is to promote the worker.

Proof In equilibrium, the firm at a given outcome, assigns the worker with a probability of one to job H if $w_H - w_L$ is strictly less than the output difference across jobs and is indifferent between jobs L and H if $w_H - w_L$ equals the output difference across jobs. Since the output difference in jobs at (i', j') strictly exceeds the output difference at (i, j) and since the worker is assigned to H with positive probability at (i, j), in equilibrium the firm must assign the worker to H with probability one at (i', j').■

Claim 2 In a skill acquisition equilibrium, the firm never promotes a worker for the outcome $(0, 0)$. Furthermore, $w_L = 0$.

Proof Suppose the firm promotes the worker with a probability of one at the outcome $(0, 0)$ in a skill acquisition equilibrium. Then from Claim 1, the only possible strategy in equilibrium is the pooling strategy (H, H, H, H). But in this case the worker has no incentive to invest in skills because his wage does not vary with the outcome, which is a contradiction. Next suppose the firm promotes the worker with a positive probability less than one. Then it must be the case that the firm is indifferent between the jobs L and H at the outcome $(0, 0)$. This implies that $w_H - w_L = 0$. Once again the worker has no incentive to acquire skills.■

Claim 3 In a skill acquisition equilibrium, the firm never promotes a worker for the outcome $(f, 0)$.

Proof Suppose the firm promotes the worker with positive probability at the outcome $(f, 0)$ in a skill acquisition equilibrium. There are then two cases to consider. First, suppose $\lambda \leq \frac{1}{2}$. Then $w_H = 0$ and since $w_L = 0$ from Claim 2, the worker has no incentive to invest in skills which is a contradiction. Next, suppose $\lambda > \frac{1}{2}$. Then there are two possibilities. If the firm does not promote the worker at $(0, g)$ then $w_H = 0$ and the worker has no incentive to invest in skills. If the firm does promote the worker with some positive probability at $(0, g)$ then $w_H = (1 - \lambda)y_H$.

In this case the firm deviates at $(0, g)$ and chooses L for sure once again leading to a contradiction.■

Claim 4 In a skill acquisition equilibrium, the firm never promotes a worker for the outcome $(0, g)$.

Proof Suppose the firm promotes the worker with positive probability at the outcome $(0, g)$ in a skill acquisition equilibrium. There are then two cases to consider. First, suppose $\lambda > \frac{1}{2}$. Then $w_H = (1 - \lambda)y_H$ and since $w_L = 0$ from Claim 2, the firm always deviates at $(0, g)$ and chooses L which is a contradiction. Next, suppose $\lambda \leq \frac{1}{2}$. If the firm does not promote the worker at $(f, 0)$, then once again $w_H = (1 - \lambda)y_H$ and the firm has an incentive to deviate and choose L at $(0, g)$. If the firm does promote the worker at $(f, 0)$ with positive probability then $w_H = 0$ and the worker has no incentive to invest in skills.■

Claim 5 In a skill acquisition equilibrium, the firm must promote a worker with positive probability for the outcome (f, g).

Proof Suppose the firm does not promote the worker at (f, g) in a skill acquisition equilibrium. The previous claims then imply that the firm's strategy is (L, L, L, L). But in this case, the worker's wage does not depend on the outcomes and he has no incentive to invest in skills which is a contradiction.■

Claim 6 In a skill acquisition equilibrium, when $\underline{p}(y_H + \Delta) < b$, the worker never specializes in a skill.

Proof Let q be the probability of the strategy (L, L, L, H) and $(1 - q)$ be the probability of (L, L, L, L). Then the only possible strategy in a skill acquisition equilibrium is one where $q > 0$. The worker's expected payoff for a given q is

$$w_L + qp_f p_g(w_H - w_L) - C(p_f, p_g)$$

First suppose $p_f = \underline{p}$ and $p_g > \underline{p}$ in equilibrium. If the worker deviates and chooses $(\underline{p}, \underline{p})$ the difference in the worker's expected payoff is

$$-q\underline{p}(p_g - \underline{p})(w_H - w_L) + C(\underline{p}, p_g) - C(\underline{p}, \underline{p})$$

Let $\underline{p}$ be sufficiently small so that $\underline{p}y_H < b$. Then

$$C(\underline{p}, p_g) - C(\underline{p}, \underline{p}) \geq (p_g - \underline{p})b > (p_g - \underline{p})\underline{p}y_H > q(p_g - \underline{p})\underline{p}(w_H - w_L)$$

where the first inequality follows from the fact that C is convex with partial derivatives bounded below by b and the last inequality follows from the fact that $w_H - w_L$ is bounded above by y_H. Thus the worker does strictly better from the deviation leading to a contradiction.

Similar reasoning can be used to rule out $p_f > \underline{p}$ and $p_g = \underline{p}$.■

Proof of Proposition 3: Consider (L, L, L, H) and $(\bar{p}, \bar{p})$. These are part of a skill acquisition equilibrium if and only if the strategies are sequentially rational for both the firm and the worker.

Notice that a worker chooses $(\bar{p}, \bar{p})$ if and only if for all (p_f, p_g) the following condition holds

$$\bar{p}^2(w_H - w_L) - (c_f + c_g)\bar{p} \geq p_f p_g(w_H - w_L) - c_f p_f - c_g p_g$$

Since the condition above must hold for all (p_f, p_g), it must hold for the following cases, i) $p_f = \underline{p}$, $p_g = \underline{p}$, ii) $p_f < \bar{p}$, $p_g = \bar{p}$, iii) $p_f = \bar{p}$, $p_g < \bar{p}$, iv) $p_f < \bar{p}$, $p_g < \bar{p}$. Substituting the values of investment in the first three cases gives us the following conditions respectively

$$w_H - w_L \geq \frac{c_f + c_g}{\underline{p} + \bar{p}} \tag{6}$$

$$w_H - w_L \geq \frac{c_f}{\bar{p}} \tag{7}$$

$$w_H - w_L \geq \frac{c_g}{\bar{p}} \tag{8}$$

For case iv) notice that because the agents expected payoff is linear in one skill when the other is fixed it follows that condition (6) is sufficient for $(\bar{p}, \bar{p})$ to be optimal.

Next, notice that wages in equilibrium must be

$$w_H = (1 - \lambda)y_H$$

$$w_L = 0$$

Also, for (L, L, L, H) to be an equilibrium strategy it must be the case that the firm does not have any incentive to deviate at each of the four outcomes $(0, 0)$, $(f, 0)$, $(0, g)$ and (f, g). These conditions are given by

$$0 \leq w_H - w_L \tag{9}$$

$$\lambda(y_H - y_L) \leq w_H - w_L \tag{10}$$

$$(1-\lambda)(y_H - y_L) \leq w_H - w_L \tag{11}$$

and

$$(y_H + \Delta) - (y_L + \Delta) \geq w_H - w_L \tag{12}$$

Substituting $w_H - w_L$ from above notice that (9) and (11) always hold. Also we can rewrite (10) as

$$\lambda \leq \frac{y_H}{(y_H - y_L) + y_H} \tag{13}$$

and (12) as

$$\lambda \geq 1 - \frac{y_H - y_L}{y_H}. \tag{14}$$

Combining (6),(7), (8), (13) and (14) gives us

$$1 - \frac{y_H - y_L}{y_H} \leq \lambda \leq min\{1 - \frac{\bar{c}}{y_H}, \frac{y_H}{(y_H - y_L) + y_H}\}$$

To show that the set of λ's satisfying the inequality is non empty notice that $\bar{c} \leq y_H - y_L$ implies $1 - \frac{y_H - y_L}{y_H} \leq 1 - \frac{\bar{c}}{y_H}$. Also notice that $1 - \frac{(y_H - y_L)}{y_H} \leq \frac{y_H}{y_H + (y_H - y_L)}$ if and only if $(y_H - y_L)^2 \geq 0$ which always holds.

Now to show that these strategies are unique, consider the firm first. For the firm, the uniqueness of (L, L, L, H) follows from Proposition 2. Next consider the worker.

We know from Proposition 2 that when $\underline{p} y_H < b$ there cannot be skill acquisition equilibria with the minimal investment level for any of the skills. We still need to rule out other cases besides $(\bar{p}, \bar{p})$ as possible equilibrium investment levels.

First, suppose $p_f = \bar{p}$ and $\underline{p} < p_g \neq \bar{p}$ in equilibrium. Since the worker's expected utility is linear with respect to p_g with p_f fixed at $\bar{p}$, it follows that $(\bar{p}, \underline{p})$ yields exactly the same expected payoff for the worker and this is given by

$$\bar{p}\underline{p}(w_H - w_L) - c_f\bar{p} - c_g\underline{p}$$

If the worker deviates and chooses $(\underline{p}, \underline{p})$ the difference in the worker's expected payoff is

$$-\underline{p}(\bar{p} - \underline{p})(w_H - w_L) + c_f(\bar{p} - \underline{p})$$

Let $\underline{p}$ be sufficiently small so that $\underline{p}y_H < b$. Then the worker can do strictly better from the deviation leading to a contradiction.

Using the same reasoning above we can rule out $p_g = \bar{p}$ and $\underline{p} < p_f \neq \bar{p}$ as an optimal choice for the worker.

Next suppose the equilibrium strategy (p_f, p_g) was interior. Since the expected payoff of the worker is linear in p_f holding p_g fixed it follows that by choosing $p_f = \underline{p}$ and p_g the worker gets exactly the same payoff. But by the reasoning above the worker can do strictly better by deviating and choosing $(\underline{p}, \underline{p})$ leading to a contradiction.■

Proof of Proposition 4: For all of the cases below we take limits as $\underline{p}$ tends to 0 and $\bar{p}$ tends to 1 when computing profits. First, consider the case where contracts cannot be written on jobs. Because $\frac{y_H}{2} > \bar{c}$ and condition (1) in Proposition 3 holds, it follows that the interval $\lambda \in (\max\{\underline{\lambda}, \frac{1}{2}\}, \bar{\lambda}]$ is non empty. The expected profit for a firm when a worker is trained in this interval is $\lambda y_H + \Delta$ and the expected profit when a worker is not trained is 0. So the returns from training are $\lambda y_H + \Delta$ when $\lambda \in (\max\{\underline{\lambda}, \frac{1}{2}\}, \bar{\lambda}]$.

Now consider the case where contracts can be written on jobs and suppose the worker is not trained in general skills so that $p_g = \underline{p}$. Then the promotion strategy (L, H, L, H), along with investment levels $(\bar{p}, \underline{p})$, contracted wages of $w_H = max\{c_f, (1-\lambda)(y_H - y_L)\}$ for job H and $w_L = 0$ for job L, and wage offers of 0 for either job are part of an equilibrium when $\lambda > \frac{1}{2}$. To see this, note that because $\lambda > \frac{1}{2}$ and $\frac{y_H}{2} > \frac{y_L}{2} + c_f$, the firm's promotion strategy is optimal at all the four outcomes. Also, the workers marginal expected utility is at least as large as $max\{c_f, (1-\lambda)(y_H - y_L)\} - c_f$ which is non-negative. Thus it is optimal for the worker to choose $p_f = \bar{p}$. The expected profit for this equilibrium is $\lambda y_H - max\{c_f, (1-\lambda)(y_H - y_L)\}$.

Finally, consider the case where contracts can be written on jobs and workers are trained in general skills. From Claim 1 in the proof of Proposition 2, there are only four possible strategies to consider for a skill acquisition equilibrium: (L, L, L, H), (L, L, H, H), (L, H, L, H) and (L, H, H, H). We will show that for any of these strategies, expected profits never exceed $\lambda y_H + \Delta$. As a result, the return from training in the interval $\lambda \in (\max\{\underline{\lambda}, \frac{1}{2}\}, \bar{\lambda}]$ is strictly larger when jobs cannot be contracted on. First, consider the strategy (L, L, L, H). Because wages in equilibrium must be at least as large as competing offers of 0 for job L, and $(1 - \lambda)y_H$ for job H, and because $\lambda y_H + \Delta > \lambda y_L > (1 - \lambda)y_L$, expected profits can never exceed $\lambda y_H + \Delta$ for this strategy. Second, consider the strategy (L, L, H, H). This strategy induces specialization in the general skill and the expected profit (as $\underline{p}$ tends to 0 in the limit) is $-(1 - p_g)w_L + p_g((1 - \lambda)y_H - w_H)$. Because wages in equilibrium must be at least as large as competing offers of 0 for job L, and $(1-\lambda)y_H$ for job H, expected profits do not exceed $\lambda y_H + \Delta$. Third, consider the strategy (L, H, L, H). This strategy induces specialization in the firm specific skill and the expected profit (as $\underline{p}$ tends to 0 in the limit) is $-(1 - p_f)w_L + p_f(\lambda y_H - w_H)$. Because wages are non-negative, expected profits can never exceed $\lambda y_H + \Delta$. Finally, consider the strategy (L, H, H, H). Notice that when $\bar{p}$ is sufficiently close to 1, workers will only specialize in one skill. To see this, suppose to the contrary that $p_f > \underline{p}$ and $p_g > \underline{p}$. Consider two possible cases. First suppose either of the investments is at the maximum level (say without loss of generality that $p_f = \bar{p}$). Then the first order necessary condition for the worker with respect to p_g is $(1 - \bar{p})(w_H - w_L) - c_g \geq 0$. Notice that since $w_H \leq y_H$ for the firm to make a positive profit and since $\bar{p}$ is sufficiently close to 1, the first order conditions do not hold, resulting in a contradiction. Alternatively, suppose (p_f, p_g) is interior. Because a worker's marginal utility is linear in either variable, it must be the case that $(\bar{p}, p_g)$ is also optimal for the worker. But once again, the worker can do strictly better by setting $p_g = \underline{p}$ which results in a contradiction. So the only possible investments in equilibrium for the strategy (L, H, H, H) are those where the worker specializes in a skill. Once again, because wages in equilibrium are non-negative expected profits cannot exceed $\lambda y_H + \Delta$.■

References

Acemoglu, D., and J.-S. Pischke (1998): "Why do Firms Train? Theory and Evidence," *Quarterly Journal of Economics*, 113, 79–119.

——— (1999): "The Structure of Wages and Investment in General Training," *Journal of Political Economy*, 107(3), 539–571.

Baker, G. (1992): "Incentive Contracts and Performance Measurement," *Journal of Political Economy*, 100, 598–614.

Baker, G., R. Gibbons, and K. Murphy (1994): "Subjective Performance Measures in Optimal Incentive Contracts," *Quarterly Journal of Economics*, 109(4), 1125–1156.

Balmaceda, F. (2005): "Firm Sponsered General Training," *Journal of Labor Economics*, 23(1), 115–133.

Becker, G. (1964): *Human Capital*. National Bureau of Economic research, New York.

Becker, G., and K. Murphy (1992): "The Division of Labor, Coordination Costs and Knowledge," *Quarterly Journal of Economics*, 107(4), 1137–1160.

Bernhardt, D. (1995): "Strategic Promotion and Compensation," *Review of Economic Studies*, 62, 315–339.

Black, S., and L. Lynch (2004): "What's driving the new economy?: the benefits of workplace innovation.," *Economic Journal*, 114(493), F97–F116.

Cappelli, P., and D. Neumark (2001): "Do High Performance Work Practices Improve Establishment-Level Outcomes?," *Industrial and Labor Relations Revew*, 54(4), 737–775.

Carmichael, L. (1983): "Firm-Specific Human Capital and Promotion Ladders," *Bell Journal of Economics*, 14(1), 251–258.

Carmichael, L., and W. MacLeod (1993): "Multiskilling, Technical Change and the Japanese Firm," *The Economic Journal*, 103(416), 142–160.

Chang, C., and Y. Wang (1995): "A Framework for Understanding Differences in Labor Turnover and Human Capital Investment," *Journal of Economic Behavior and Organization*, 28(1), 91–105.

——— (1996): "Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited," *Journal of Labor Economics*, 14(3), 505–519.

DeVaro, J., S. Ghosh, and C. Zoghi (2008): "Job Characteristics and Labor Market Discrimination in Promotions: New Theory and Empirical Evidence," *Working Paper*, pp. 1–54.

DeVaro, J., and M. Waldman (2011): "The signaling role of promotions: Further theory and empirical evidence," *Journal of Labor Economics (forthcoming)*.

Gant, J., C. Ichniowski, and K. Shaw (2002): "Social Capital and Organizational Change in High-Involvement and Traditional Work Organizations," *Journal of Economics and Management Strategy*, 11(2), 289–328.

GHOSH, S., AND M. WALDMAN (2010): "Standard Promotion Practices versus Up-or-out Contracts," *Rand Journal of Economics*, 41, 301–325.

GIBBONS, R., AND L. KATZ (1991): "Layoffs and Lemons," *Journal of Labor Economics*, 9(4), 351–380.

GIBBS, M., AND A. LEVENSON (2002): *The Expansion of Economics, (S. Grossbard-Schechtman and C. Clague Eds.)*chap. The Economic Approach to Personnel Research. M.E Sharpe, Inc., New York.

HOLMSTROM, B., AND P. MILGROM (1991): "Multitask Principal-Agent Analysis: Incentive Contracts, Job Ownership and Asset Design," *Journal of Law, Economics and Organization*, 7, 24–52.

ICHNIOWSKI, C., AND K. SHAW (2003): "Beyond Incentive Pay: Insiders Estimates of the Value of Complementary Human Resource Management Practices," *Journal of Economic Perspectives*, 17, 155–178.

ICHNIOWSKI, C., K. SHAW, AND G. PRENNUSHI (1997): "The Effects of Human Resource Management Practices on Productivity," *American Economic Review*, 86, 291–313.

KAHN, C., AND G. HUBERMAN (1988): "Two-Sided Uncertainty and "Up-or-Out Contracts"," *Journal of Labor Economics*, 6(4), 423–444.

KATZ, E., AND A. ZIDERMAN (1990): "Investment in General Training: The Role of Information and Labour Mobility," *Economic Journal*, 100(403), 1147–1158.

KESSLER, A., AND C. LULFESMANN (2006): "The Theory of Human Capital Revisited: On the Interaction of Specific and General Investments," *Economic Journal*, 116(514), 903–923.

LAWLER, E., S. MOHRMAN, AND G. LEDFORD (1995): *Creating High Performance Organizations: Practices and Results of Employee Involvement and TQM in Fortune 1000 Companies.* Jossey Bass, San Francisco.

LEVIN, J. (2003): "Relational Incentive Contracts," *American Economic Review*, 93(3), 835–857.

LINDBECK, A., AND D. SNOWER (2000): "Multitask Learning and Reorganization of Work: From Tayloristic to Holistic Organization," *Journal of Labor Economics*, 18(3), 353–376.

MACDONALD, G., AND L. MARX (2001): "Adverse Specialization," *Journal of Political Economy*, 109(4), 864–899.

MACLEOD, W., AND J. MALCOMSON (1988): "Reputation and Hierarchy in Dynamic Models of Employment," *Journal of Political Economy*, 96(4), 832–854.

MALCOMSON, J. (1984): "Work Incentives, Hierarchy, and Internal Labor Markets," *Journal of Political Economy*, 92(3), 486–507.

MALCOMSON, J., J. MAW, AND B. MCCORMICK (2003): "General Training by Firms, Apprentice Contracts, and Public Policy," *European Economic Review*, 47(2), 197–227.

MILGROM, P., AND S. OSTER (1987): "Job Discrimination, Market Forces, and the Invisibility Hypothesis," *Quarterly Journal of Economics*, 102(3), 453–476.

MYERSON, R. (1978): "Refinement of the Nash Equilibrium Concept," *International Journal of Game Theory*, 7, 73–80.

OSTERMAN, P. (1994): "How Common is Workplace Transformation and Who Adopts it?," *Industrial and Labor Relations Revew*, 47(2), 173–188.

——— (2000): "Work Reorganization in an Era of Restructuring: Trends in Diffusion and Effects on Employee Welfare.," *Industrial and Labor Relations Revew*, 53(2), 179–196.

O'TOOLE, J., AND E. E. LAWLER (2006): *The New American Workplace.* Palgrave-Macmillan.

OWAN, H. (2004): "Promotion, Turnover, Earnings and Firm-Sponsored Training," *Journal of Labor Economics*, 22(4), 955–978.

PRENDERGAST, C. (1993): "The Role of Promotions in Inducing Specific Human Capital Acquisition," *Quarterly Journal of Economics*, 108(2), 523–534.

ROSEN, S. (1978): "Substitution and Division of Labor," *Economica*, 45(179), 235–250.

ROY, A. (1950): "The Distribution of Earnings and of Individual Output," *Economic Journal*, 60, 489505.

SCHONBERG, U. (2007): "Testing for Asymmetric Employer Learning," *Journal of Labor Economics*, 25(4), 651–691.

SCOONES, D., AND D. BERNHARDT (1998): "Promotion, Turnover and Discretionary Human Capital," *Journal of Labor Economics*, 16(1), 122–141.

SMITH, A. (1937): *An Inquiry into the Nature and Causes of the Wealth of Nations.* Modern Library, New York.

STEVENS, M. (1994): "A Theoretical Model of On-the-Job Training with Imperfect Competition," *Oxford Economic Papers*, 46(0), 537–562.

WALDMAN, M. (1984): "Job Assignments, Signalling, and Efficiency," *The RAND Journal of Economics*, 15(2), 255–267.

——— (1990): "Up-or-Out Contracts: A Signaling Perspective," *Journal of Labor Economics*, 8(2), 230–250.

ZABOJNIK, J., AND D. BERNHARDT (2001): "Corporate Tournaments, Human Capital Acquisition and the Firm Size-Wage Relation," *Review of Economic Studies*, 68, 693–716.