

On the Information Gathering Role of Firm-Sponsored Training for New Hires

by

Doyoung Kim
Department of Economics
Sogang University
Seoul 121-742, Korea
dkim@sogang.ac.kr

Jacques Lawarrée
Department of Economics
University of Washington
Seattle, WA 98195-3330, U.S.A.
and
ECARES
Brussels, Belgium
lawarree@u.washington.edu

October 2008

According to Becker's human capital theory, employers tend to under-invest in general training relative to specific training as it increases their existing employees' outside opportunities. We show that this is not necessarily true if training has an information gathering function that allows an agent to learn his skills. An example is the training of new hires. When training creates an information asymmetry between the principal and the agent, the principal may over-invest in general training relative to specific training. General training helps the principal reduce the incentive problem inside the firm. Becker's result of under-investment in general training may not hold when training creates asymmetric information. (JEL: D82, D83, J31, L22)

1. Introduction

In many situations, an agent who engages in certain activities sponsored by a principal can gather critical information as a byproduct. An example is the training for new hires. As new hires are often inexperienced, they may not know their abilities or skills yet. By participating in a training program sponsored by the employer (principal), newly-hired workers (agents) can learn their abilities or skills.¹ If this information is revealed to the worker only, an incentive problem arises due to the information asymmetry. In this paper, we examine how a principal decides his investment in training taking into account this incentive problem. In particular, we compare the principal's willingness to invest in general versus specific training.

It is a well-known result in the literature (BECKER [1964]) that firms are generally reluctant to invest in general training. The reason is that general training increases the human capital useful outside the firm and a newly-trained agent is able to promptly leverage the higher outside opportunities provided by general training. However, contrary to this view, recent empirical studies find evidence that firms do sponsor general training to a broad extent, particularly for young workers, new hires, and apprentices. For instance, Germany has a well-documented institutionalized apprenticeship system where workers receive a training that is largely general and whose cost is mostly paid by employers (HARHOFF AND KANE [1997] and ACEMOGLU AND PISCHKE [1998]). Similar evidence exists in other countries: e.g., see BOOTH AND BRYAN [2007] for the U.K. or GERFIN [2004] for Switzerland. In the U.S., LOWENSTEIN AND SPLETZER [1998] also report that employers mostly pay for general training without passing on the cost to workers in the form of lower wages. The BLS' National Compensation survey of 2007 shows that assistance for general education is provided by 15% of firms (23% for firms of more than 100 employees).

This evidence has prompted a renewed theoretical interest in Becker's work. For firm-sponsored general training to be profitable, theoretical studies point to the necessity of labor market frictions and suggest several sources of frictions including information asymmetry between incumbent and potential employers in addition to imperfect labor market competition.² Synthesizing these studies, ACEMOGLU AND PISCHKE [1999] conclude that wage compression associated with labor market frictions allows firms to sponsor general training. More recently, KESSLER AND LÜLFESMANN [2006] show that even without labor market frictions general training can be profitable due to an incentive complementarity with specific training.

In this paper, we do not address the issue of how general training is profitable. We just take ACEMOGLU AND PISCHKE's [1999] results as given and simply assume that general training can be profitable for the principal due to labor market frictions. We then investigate the conditions under which general training can be more profitable for the principal than specific training. At first glance, such situation cannot happen. The reason

¹ There are many other examples of this situation: for instance, when a firm invests in R&D, researchers acquire information about a new technology; or when a seller provides free samples, seminars and product demonstrations, buyers learn their preference over the product.

² See KATZ AND ZIDERMAN [1990], CHAND AND WANG [1996], ACEMOGLU AND PISCHKE [1998] for asymmetric information between incumbent and potential employers, and STEVENS [1994] and ACEMOGLU [1997] for imperfect labor market competition.

is that, unless labor market frictions are absolute, general training raises the agent's outside opportunities, which is not the case for specific training.

However, we show that the lower willingness to invest in general training relative to specific training does not necessarily hold true if training has an information gathering role and if the gathered information about the agent's skills is private to him and therefore creates an asymmetry of information. Specific training allows the agent to learn his specific skills. Similarly, general training allows the agent to learn his general skills. In addition, as general skills determine the agent's outside opportunities, general training also allows the agent to learn his outside opportunities. This information gathering role of training makes training beneficial for the principal because it allows for better matching between the workers and their job or work load. The key assumption that differentiates our model from others is the information asymmetry between the principal and the agent. If the agent privately learns his skills, training can be detrimental for the principal because the information asymmetry creates an incentive problem, hindering efficient matching. Thus, the relative profitability between general and specific training depends on which training helps the principal to reduce the incentive problem.

Regardless of whether it is general or specific skills, traditional incentive theory (e.g., LAFFONT AND TIROLE [1993]) has stressed the agent's incentive to under-report his skills, which forces the principal to give the agent an information rent. This is the incentive problem the principal faces when investing in training. We show that the private information about outside opportunities, which is only relevant for general training, actually creates a force that counteracts the agent's incentive to under-report because an agent who under-reports his general skills reduces his outside opportunities. If an agent with high general skills claims low general skills to obtain a rent from the principal, he will forgo the benefit of high outside opportunities. This does not happen with specific skills since they do not affect outside opportunities. An agent has therefore less incentive to under-report his general skills compared to specific skills, and the principal can reduce the agent's information rent in the case of general training. Furthermore, we show that the amount of the reduction in information rent depends on the gap between the outside opportunities available to high skill agents and low skill agents, as identified by general training.

Although the principal benefits from the reduction in information rents, there are costs associated with general training. General training raises the agent's general human capital and the principal must compensate the agent for his higher outside opportunities from general training. Facing this trade-off, we show that the principal over-invests in general training relative to specific training if an improvement in the value of the agent's outside opportunities is small relative to its gap between high and low skill agents.

A straightforward empirical implication of our result is that we should expect to see more general training for new hires as it has an important information gathering role. There is indeed ample evidence that this is the case. For example, the German apprenticeship fits our model well. As ACEMOGLU AND PISCHKE [2000, p.926] explain: "In the case of Germany, however, where workers enter apprenticeships at a relatively young age, like 15 or 16, it is unlikely that they have very good knowledge about all their own aptitudes and comparative advantages." LYNCH [1992] studies the training for young workers and finds that the probability of investing in general training is lower if the youth has a longer tenure on the job. The main form of general training given by the

U.S. Army forces is through the “Tuition Assistance Program” that allows new recruits to go to college (BUDDIN AND KAPUR [2005]).

Our model differs from recent models of general training in several ways. First, while we consider information asymmetry between the principal (employer) and the agent (worker), most of the literature (e.g., KATZ AND ZIDERMAN [1990] CHANG AND WANG [1996], and ACEMOGLU AND PISCHKE [1998]) consider information asymmetry between incumbent and potential employers. Second, we consider both general and specific skills whereas many models focus only on general skills. With these critical differences, our paper is able to provide an additional justification for general training by emphasizing the information gathering role of training for new hires. Furthermore, our analysis complements the prior studies by examining the relative profitability of general training focusing on frictions inside the firm (the agent’s incentive problem due to the acquisition of asymmetric information) in addition to frictions outside the firm.

Our model is closely related to KESSLER AND LÜLFESMANN [2006] in that both consider general and specific trainings and adopt contractual approaches to reconcile firm-sponsored general training. Their paper uncovers a complementarity between general and specific training investments as a result of the bargaining between the employer and employee after training. There are two main differences with our paper. First, we consider investment in general and specific training separately and therefore we do not incorporate the complementarity between the two types of training. Second, instead of bargaining, we consider a screening contract that sort agents into their levels of skills. With these differences, our model generates different empirical implications. For instance, KESSLER AND LÜLFESMANN [2006] show that the presence of specific skills is necessary for investment in general training, so general training is more likely for workers with several years of tenure (i.e., workers with accumulated specific skills). In contrast, our model suggests that general training is more likely for workers with little tenure such as new hires and apprentices. For cross-firm implications, KESSLER AND LÜLFESSMAN [2006] argue that general training is more likely to be observed in large firms because more specific skills are put in place in such firms. While our model may not generate an implication about firm size effects, it indicates that general training is more likely to be observed in firms with diverse workers’ general abilities because the reduction in information rents from training is larger for such firms.

Furthermore, our model offers a testable prediction about the impact of training on skill premiums associated with asymmetric information. A standard adverse selection model suggests that asymmetric information allows high skill agents to command rents, implying that skill premiums become greater under asymmetric information. Our model shows that general training creates a force countering high skill agents’ incentive, which in turn reduces rents given to them. This implies that in the presence of asymmetric information, skill premiums are small for firms investing in general training relative to firms investing in specific training.

When comparing general training with specific training, LEWIS AND SAPPINGTON [1993] consider asymmetric information inside the firm like we do. They examine a model where outside opportunities affect the training effort by the agent. Versatile workers have private information about their outside opportunities that determine their wages, and have an incentive to overstate their outside opportunities. As the optimal output is distorted upward in the presence of such a countervailing incentive, training for

versatile workers (general training) that increases the output adds another upward-distortion. General training becomes less attractive for the principal. The main difference between their model and ours is that training in their model is just an investment in human capital, not an information gathering activity as the agent already knows his outside opportunities before training. This results in under-investment in general training relative to specific training.

While we focus on the application of our model to the training context, our paper also contributes to the literature on information gathering. We extend the literature by studying how information about the agent's outside opportunities affects the incentive of the firm to invest in information gathering.³ The literature on information gathering is recent and still developing. In CRÉMER AND KHALIL [1992], the agent's information gathering is strategic,⁴ which makes the principal not invest in information gathering. LEWIS AND SAPPINGTON [1997], and CRÉMER, KHALIL, AND ROCHET [1998a] consider productive information gathering like we do. However, LEWIS AND SAPPINGTON [1997] assume that the principal always prefer information gathering to no information, which makes an analysis of investment in information gathering irrelevant. CRÉMER, KHALIL, AND ROCHET [1998a] analyze information gathering in terms of the agent's information gathering cost. However, unlike us, they do not consider the case where information gathering reveals the agent's outside opportunities. In the literature, it is usually the agent who decides whether to gather information.⁵ An exception is LEWIS AND SAPPINGTON [1994]. In their model, the principal can endow the agent with private information similar to us. However, information gathering is costless, and it does not affect the agent's outside opportunities.

Our paper is organized as follows. In section 2, we present the basic model of training with information gathering. In section 3, we analyze the effect of the identification of the agent's skills and outside opportunities on the level of investment in training. In section 4, we briefly discuss some extensions to the basic set-up. It includes the cases where general and specific skills are correlated with each other, where the agent pays training costs, and where training affects the agent's skills differently across the agent's types. We also introduce another application of our model to the context of R&D investment. Section 5 concludes the paper.

2. Model

Skills and outside opportunities

In a principal-agent framework, we present a model focusing on investment in training for new hires.⁶ A principal hires an agent to carry out the production of a good ($q \geq 0$) and gives him a transfer t for production. The value of the product for the

³ Our model is also built upon the literature on countervailing incentives as the acquisition of private information about outside opportunities creates such incentives. For countervailing incentives, see LEWIS AND SAPPINGTON [1989a], [1989b], LAFFONT AND TIROLE [1989], JULLIEN [2000], and BONTEMS AND BOURGEON [2000]. This literature, however, does not consider the incentive to invest in information gathering.

⁴ See also CRÉMER, KHALIL, AND ROCHET [1998b] for strategic information gathering.

⁵ See also GEHRIG [2004] for the role of various organizational forms on information acquisition incentives.

⁶ Our model can be applied to other frameworks involving information gathering. See section 4.4 for an application of our model to R&D investment.

principal is given by $V(q)$, where $V(\cdot)$ is increasing, concave, twice differentiable and satisfies the Inada conditions ($V'(0) = \infty$ and $V'(\infty) = 0$). The cost of production is borne by the agent and it is $(\beta^G + \beta^S)q$ for producing q , with β^G and β^S being the cost parameters. We interpret the cost parameters as the skills of the agent. Every agent has two types of skills: general and specific. Accordingly, β^G and β^S capture general and specific skills, respectively. General skills can be productively used not only for the current employer but also for other employers. Specific skills, on the other hand, are only useful to the current employer. An agent with low cost parameters is a highly skilled or highly productive agent who incurs a small marginal cost of production.

At the outset, neither the principal nor the agent knows the level of the agent's skills. Instead, they simply share the common belief that (i) β^G is β_H^G with probability π_H^G or β_L^G with probability π_L^G ($\pi_H^G + \pi_L^G = 1$, $\Delta\beta^G \equiv \beta_L^G - \beta_H^G > 0$); (ii) β^S is β_H^S with probability π_H^S or β_L^S with probability π_L^S ($\pi_H^S + \pi_L^S = 1$, $\Delta\beta^S \equiv \beta_L^S - \beta_H^S > 0$). For simplicity, we assume identical distributions such that $\pi_i \equiv \pi_i^G = \pi_i^S$, $i = H, L$. We denote the (ex-ante) expected level of skills by $\tilde{\beta}^G \equiv \pi_H \beta_H^G + \pi_L \beta_L^G$ and $\tilde{\beta}^S \equiv \pi_H \beta_H^S + \pi_L \beta_L^S$. In our base model, we assume that the skills are independent. Our model is extended to the case of a positive correlation between general and specific skills in section 4.1.

The value of the agent's outside opportunities U is determined by his general skills so that an agent with high general skills (β_H^G) has the high value of outside opportunities (U_H). Accordingly, as are general skills, the value of outside opportunities is unknown at the outset. The principal and the agent share the common belief that U is U_H with probability π_H or U_L with probability π_L ($\Delta U \equiv U_H - U_L > 0$), where the (ex-ante) expected value of outside opportunities is given by $\tilde{U} \equiv \pi_H U_H + \pi_L U_L$. In contrast to general skills, the agent's specific skills are only indirectly related to the value of outside opportunities through their relationship with general skills. The assumption of zero correlation between general and specific skills in the base model implies that both agents with high and low specific skills have the value of outside opportunities equal to \tilde{U} . ***Training Program***

The principal sets up a training program for the agent. Before engaging in production, the agent participates in the training program. Training can be directed toward the agent's general skills (it is then called general training) or toward his specific skills (it is then called specific training), but not both. This distinction between the two types of training goes back to BECKER [1964]. We recognize that it is not always possible to distinguish between them since real world training often affects both types of skills simultaneously.⁷ However, our simplifying assumption allows us to sharply contrast between general and specific trainings by separating the incentive effects of each type of training.

Training for newly hired agents has an information gathering function. It allows

⁷ Sometimes the distinction can be quite clear. Pharmaceutical representatives generally attend motivational seminars to improve their sales abilities (general training) as well as lessons about the particular drugs sold by their company (specific training).

the agent to identify his skills. The recent literature has stressed the role of skill identification in training. It is, for instance, a central theme in AUTOR [2001]. Unlike training for senior agents who may already know their skills, training for newly hired junior agents provides them with an opportunity to learn their skills, productivity or ability. We do not posit that without training, the agent would never identify his skills. The agent may learn his skills over the course of his tenure in the firm. But our model is relevant if training expedites a learning process that would otherwise take years and may require the agent to move between many positions. The information gathering role of training is essential since firms can then allocate tasks or work load efficiently according to the agents' skills. This identification process occurs in specific and general training. That is, general training allows the agent to identify his general skills (β_H^G or β_L^G) and specific training allows the agent to identify his specific skills (β_H^S or β_L^S). The critical difference is that the agent can also identify his outside opportunities (U_H or U_L) from general training as general skills determine the agent's outside opportunities.⁸ The information gathered is private to the agent, so that no other parties know the agent's skills. If the information gathered is common knowledge, there is no incentive problem and the analysis is straightforward (see appendix).

While the agent can identify the potential value of his outside opportunities from general training, he may not be able to collect this value when he moves to another firm because of labor market frictions. Such frictions may be due to an information asymmetry between the principal and other potential employers when potential employers are not aware of the general training.⁹ For instance, if potential employers do not observe whether general training took place,¹⁰ they would offer a pooling wage contract to the agent based on \tilde{U} .¹¹ In contrast, if potential employers do know that general training took place (but they still do not know the agent's type as it is private

⁸ Our model assumes that the agent learns his skills perfectly. Alternatively, we can allow for stochastic learning with the agent learning his skills only with some probability as in KESSLER [1998]. As long as the learning process is the same for general and specific trainings, stochastic learning would not qualitatively change our results that compare the two trainings.

⁹ As ACEMOGLU AND PISCHKE [1999] point out, the presence of labor market frictions is necessary for general training to be profitable for the principal. Note that our focus is not on the profitability of general training per se, which is however necessary for us to investigate the relative profitability of general training to specific training. Formally, one may introduce the agent's turnover process and his interaction with potential employers to show how labor market frictions allow for profitable general training. However, as we highlight the relative profitability of general training hinged on the frictions (information asymmetry) inside the firm, the formal modeling of the frictions outside the firm simply adds complexity to our model without generating new insights. Accordingly, we just assume labor market frictions similar to ACEMOGLU AND PISCHKE [1998] and let their results built in our model: labor market frictions prevent potential employers from discerning the identity of an agent who left the incumbent employer and therefore make them offer a pooling contract.

¹⁰ If an employee can produce a certificate after undergoing general training, a high skill agent has an incentive to reveal this certificate to the potential employers to separate himself from low skill agents or untrained agents, and therefore the potential employer would know that such training took place. However, current employers have no incentive to provide such a certificate as it would only help the agent leverage his new outside opportunities (see ACEMOGLU AND PISCHKE [2000] and HECKMAN, ROSELIUS, AND SMITH [1994]).

¹¹ Due to the unobservability, potential employers do not know whether or not the agent identifies the value of outside opportunities. Formally, this is equivalent to the case where the principal (potential employers) faces possible ignorant agents. As LEWIS AND SAPPINGTON [1993] show, pooling occurs in equilibrium.

information), it becomes common knowledge that the agent has identified the value of his outside opportunities, U_H or U_L , and a screening contract separating the high and low skill agents is optimal.

To take such frictions into account in a simple way, the value of the agent's outside opportunities he can collect after general training is formalized as $\gamma U_i + (1-\gamma)\tilde{U}$, where $\gamma \in (0,1)$ and $i = H, L$. We can interpret $(1-\gamma)$ as the degree of labor market frictions hindering the relationship between the true potential value of the agent's outside opportunities and the collectible value. While market frictions prevent the agent from collecting the full value of his outside opportunities from potential employers, our focus is on the information revelation inside the firm. One of our contributions is to analyze the asymmetry of the information acquired in the identification process of training between the principal and the agent. We assume that only the agent finds out the true value of his skills as it is the agent, not the principal, who participates in the training program.¹² This asymmetry turns out to be critical in determining the principal's incentive to invest in specific and general training. To be clear, it should be noted that neither the principal nor the potential employers have information about the agent's type. However, the principal knows whether training takes place, while the potential employers only know it with probability γ . In this respect, the potential employers are less informed than the incumbent employer for all $\gamma < 1$.

In addition to the information gathering role of training, we also follow the standard assumption that training improves the agent's skills. We denote by α^G and α^S the improvement in general and specific skills due to training. More specifically, an agent of type β_i^G (β_i^S) sees his production costs reduced to $\theta_i^G \equiv \beta_i^G - \alpha^G$ ($\theta_i^S \equiv \beta_i^S - \alpha^S$) after training. To focus on the effects of the identification of innate skills (β_i^G and β_i^S), we assume that α^G and α^S are constant over the agent's type: $\alpha_H^G = \alpha_L^G = \alpha^G$ and $\alpha_H^S = \alpha_L^S = \alpha^S$.¹³ Unlike AUTOR [2001], we do not need to assume that training is more productive in a marginal sense for agents with high skills. What is required in our model is that an agent with high innate skills (β_H^G or β_H^S) is still a high skill agent after training ($\theta_H^G > \theta_L^G$ and $\theta_H^S > \theta_L^S$). In the case of general training, the value of the agent's outside opportunities increases by K thanks to the accumulation of general human capital.¹⁴ Again, due to labor market frictions, the agent is only able to collect γK .

We assume that the principal bears training costs C up front. We discuss the case where the agent pays for training later in section 4.2. If the principal does not invest in training, the agent produces without knowing the cost parameters, and thus information

¹² There is another asymmetric information case where only the principal finds out. However, this informed principal case is not very interesting. Since the principal's payoff does not depend directly on the agent's skills, there is no screening condition for the principal. In other words, there is no profitable mechanism that induces a truthful revelation from the principal (See MASKIN AND TIROLE [1990], [1992]).

¹³ As will be clear later, since α^G and α^S are inside efficiency parameters, there will be no change in our main result regarding the effect of outside opportunities on investment in training even if we allow for $\alpha_H^G \neq \alpha_L^G$ and $\alpha_H^S \neq \alpha_L^S$. See section 4.3 for further discussion.

¹⁴ In our base model, K is type-independent ($K_H = K_L = K$). We relax the assumption of constant K in section 4.3.

from the training is productive in the sense of CRÉMER, KHALIL, AND ROCHET [1998a]. Note that since the cost parameters, β^G and β^S , represent the agent's skills acquired before the contract, his ex-post utility depends on the true values of β^G and β^S no matter whether information gathering takes place or not.¹⁵ We assume that the agent has the option of terminating his relationship with the principal at any time. Technically, this assumption compels the principal to take account of the effect of information gathering on the agent's outside opportunities.

Timing

The timing of the game is summarized as follows. At the outset, no one (the principal, the agent, and potential employers) knows the agent's skills, but they share a common belief about β^G , β^S , and U . The principal offers the agent a contract specifying the transfer t contingent on the output q and a training program for the agent. General training reveals that the agent of expected type $\tilde{\beta}^G$ is either type θ_H^G or θ_L^G . Similarly, specific training reveals that the agent of expected type $\tilde{\beta}^S$ is either type θ_H^S or θ_L^S . If training is general, the agent can expect his outside opportunities to be valued either at $\gamma U_H + (1-\gamma)\tilde{U} + \gamma K$ or at $\gamma U_L + (1-\gamma)\tilde{U} + \gamma K$ depending on his type. When there is specific training or no training, the value of the agent's outside opportunities is simply \tilde{U} . Finally, the agent engages in production. The output q and the transfer t are realized thereafter.¹⁶

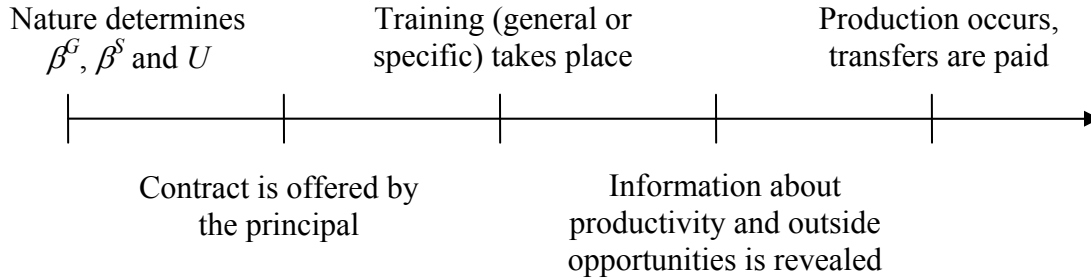


Figure 1: The timing of the game

Benchmark contracts

The principal will invest in training only if it is profitable to do so and therefore compare her objective function with and without training. We first present here the

¹⁵ Therefore, the agent's true cost will eventually be revealed to him, but this information will come too late to affect the contract. This is a standard assumption in the information gathering literature.

¹⁶ Notice that standard training models including BECKER [1964] consider a two-period production: in the first period the agent participates in a training program and at the same time engages in production, and in the second period only production takes place. This two-period production model allows an agent with credit constraints to pay for training by accepting lower wage for his first-period production. Since we take the profitability of firm sponsored training as given, we simplify the model by eliminating the first-period production.

optimal contract without training. When the principal does not invest in training, the agent produces the output and receives a monetary transfer based on the expected level of skills $(\tilde{\beta}^G + \tilde{\beta}^S)$ and the expected value of outside opportunities (\tilde{U}) . The contract that maximizes the principal's objective function $(V(q)-t)$ subject to the agent's participation constraint $(t - (\tilde{\beta}^G + \tilde{\beta}^S)q \geq \tilde{U})$ gives the optimal output, denoted by \tilde{q} , satisfying the first order condition $V'(\tilde{q}) = \tilde{\beta}^G + \tilde{\beta}^S$. At the optimum, the value of the objective function is

$$W^{NT} = V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S)\tilde{q} - \tilde{U}. \quad (1)$$

While relegating the details to the appendix, as a benchmark, we briefly discuss the principal's incentive to invest in general training relative to specific training when training produces symmetric information (both the principal and the agent learn the agent's skills). To make a comparison between general and specific trainings, we consider a case where $\beta_i^G = \beta_i^S$ and $\alpha^G = \alpha^S$ for $i = H, L$. The contract is now based on the true level of skills. The value of the agent's outside opportunities is also changed depending on whether training is directed toward general skills or specific skills. As training, however, produces "symmetric" information, not surprisingly, the optimal output is the same for specific and general trainings. Given that both trainings improve the inside productivity to the same amount ($\alpha^G = \alpha^S$), there is no reason to see the difference in the optimal output. However, for the same output produced by the agent, the principal has to pay additional compensation to the agent to make up for γK in case of general training. It makes the principal under-invest in general training relative to specific training.¹⁷ In the extreme case where there are no labor market frictions ($\gamma = 1$), the agent is able to extract all the outside benefits of general training (K), and the principal may not invest at all in general training.¹⁸

3. Investment in general training relative to specific training

The result of under-investment in general training relative to specific training may cease to be true if training produces asymmetric information, as we shall see in this section. In what follows, we first characterize investment in specific training as the solution to a standard adverse selection problem. Next, we derive the solution to the investment in general training and show the effect of outside opportunities as countervailing incentives. We then compare the levels of investment in specific and general training by fully characterizing the solution to the investment in general training as a function of the gap between the outside opportunities of high and low skill agents.

¹⁷ Note that there are, at least, two well-known solutions to the problem of under-investment in general training. The first one is to let the workers pay for general training. We consider this case in section 4.2. However, training costs can be substantial and credit constrained workers may not be able to pay those costs up front. Another solution is to let the firm recoup the training costs by paying the training worker less than his improved post-training marginal product in a long-term contract. In this model, we assume that firms and workers cannot commit to long-term wage contracts.

¹⁸ This corresponds to Becker's result that employers are less willing to pay for training that enhances outside opportunities for the workers.

Investment in specific training

With specific training, the agent learns his specific skills, but not general skills or the value of outside opportunities. The contract is based on the true level of specific skills ($\theta_i^S \equiv \beta_i^S - \alpha^S$), the expected level of general skills ($\tilde{\beta}^G$), and the expected value of outside opportunities (\tilde{U}). The principal's problem is the standard second best problem under asymmetric information:

$$\begin{aligned} & \text{Max} \sum_{q_i, t_i} \pi_i [V(q_i) - t_i] - C \\ \text{s.t. } & (IR_i) \quad t_i - (\tilde{\beta}^G + \theta_i^S)q_i \geq \tilde{U} \quad \forall i \in \{H, L\} \\ & (IC_i) \quad t_i - (\tilde{\beta}^G + \theta_i^S)q_i \geq t_j - (\tilde{\beta}^G + \theta_j^S)q_j \quad \forall i, j \in \{H, L\}. \end{aligned}$$

Asymmetric information generates the traditional incentive for the agent with high specific skills to pretend to be a low skill agent in order to be assigned less arduous work or output. To prevent misrepresenting, the agent must be given an information rent.

The adverse selection created by training makes it uncertain whether training is a profitable investment for the principal. The trade-off is between efficiency gains from both a reduction of marginal cost and a better match between each type and the output on the one hand, and a positive rent given to the agent with high specific-skills on the other hand. Formally, the principal will invest in information gathering if the cost of information gathering C is less than $C^S = W^S - W^{NT}$, where W^S is the value of the objective function gross of C in this problem. The principal's maximum willingness to pay for specific training, C^S , can be derived explicitly:

$$C^S = \left\{ \sum_{i \in \{H, L\}} \pi_i [V(q_i^S) - (\tilde{\beta}^G + \theta_i^S)q_i^S] \right\} - [V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S)\tilde{q}] - \pi_H \Delta\beta^S q_L^S. \quad (2)$$

where q_i^S is defined as $V'(q_H^S) = \tilde{\beta}^G + \theta_H^S$ and $V'(q_L^S) = \tilde{\beta}^G + \theta_L^S + \frac{\pi_H}{\pi_L} \Delta\beta^S$. The first two terms of the above expression represent the efficiency gains from both a reduction of marginal cost and a better match while the last term is the rent to the high skill agent.¹⁹

Investment in general training

With general training, instead of expecting \tilde{U} as outside opportunity, the agent knows that his outside opportunities are valued at $\gamma U_H + (1-\gamma)\tilde{U} + \gamma K$ if he is an agent with high general skills and $\gamma U_L + (1-\gamma)\tilde{U} + \gamma K$ if he is an agent with low general skills. The principal's problem can be stated as

¹⁹ We follow a typical assumption in the literature on information gathering that there is an efficiency gain from identification, i.e., even if $\alpha^S = 0$ and thereby $\theta_i^S = \beta_i^S$, $C^S > 0$ (see, for example, LEWIS AND SAPPINGTON [1997]).

$$\begin{aligned}
& \text{Max}_{q_i, t_i} \sum_{i \in \{H, L\}} \pi_i [V(q_i) - t_i] - C \\
& \text{s.t. } (IR_i) \quad t_i - (\theta_i^G + \tilde{\beta}^S)q_i \geq \gamma U_i + (1 - \gamma)\tilde{U} + \gamma K \quad \forall i \in \{H, L\} \\
& \quad (IC_i) \quad t_i - (\theta_i^G + \tilde{\beta}^S)q_i \geq t_j - (\theta_j^G + \tilde{\beta}^S)q_j \quad \forall i, j \in \{H, L\}.
\end{aligned}$$

Once he has identified higher outside opportunities, the agent with high general skills can now expect a higher transfer. This discourages him from pretending to be an agent with low general skills. In other words, the identification of outside opportunities generates a force that counteracts the traditional incentive for the high skill agent,²⁰ and allows the principal to ease the output distortion. However, if the gap between the outside opportunities of the high and low skill agents (ΔU) is large, the agent with low general skills develops an incentive to pretend to be an agent with high general skills. The principal gives a rent to the agent with low general skills and distorts the output for the agent with high general skills upward, in order to dissuade the agent with low general skills from pretending to be an agent with high general skills.

As the optimal contract varies with ΔU ,²¹ the principal's maximum willingness to pay for general training, denoted by C^G , also varies with ΔU . We compute this formally in the appendix.

General training vs. specific training

We are now in a position to analyze the principal's incentive to invest in general training relative to specific training. Again, to make a comparison between general and specific trainings, we consider a case where $\beta_i^G = \beta_i^S \equiv \beta_i$ and $\alpha^G = \alpha^S \equiv \alpha$ for $i = H, L$. That is, a priori general and specific skills have the same effects on the inside productivity. The question is then whether C^G is greater or less than C^S in (2). The answer depends on the gap between the outside opportunities of the high and low general skill agents (ΔU) relative to the improvement in the value of the agent's outside opportunities from general training (K) as summarized in the following proposition.

Proposition 1 *For $\gamma \in (0, 1)$ and $\Delta U < \overline{\Delta U} \equiv \frac{1}{\gamma} \Delta \beta q_H^*$, the principal always over-invests*

²⁰ There are other forces identified in the literature such as career concerns and promotion that can counteract the agent's traditional incentive to shirk. In a moral hazard framework, DEWATRIPONT, JEWITT AND TIROLE [1999], who generalize the original contribution by HOLMSTRÖM [1982], show in the presence of career concerns that the agent would exert effort or take arduous tasks even without an explicit incentive. Similarly, LAZEAR AND ROSEN [1981] show that a tournament competition for promotion makes the agent exert first-best effort. Formally, career concerns and promotion have similar effects on the agent's incentive as general training because all of them affect the RHS of the individual rationality constraint, which defines the agent's current outside opportunities, future expected opportunities, job tenure or option values depending on the modeling choice. If we model both adverse selection and career concerns for instance, the agent's incentive may not be uniquely characterized in equilibrium as signaling becomes a relevant issue. Nevertheless, unless career concerns nullify the traditional incentive, general training helps counteract the traditional incentive.

²¹ See LEWIS AND SAPPINGTON [1989a] and JULLIEN [2000] for an optimal contract with countervailing incentives in general.

in general training relative to specific training for small values of K ($K < \bar{K} \equiv \pi_H \Delta U$). For larger values of K , over-investment may still occur as long as K is not too large ($K < \bar{\bar{K}}$, where $\bar{\bar{K}} > \bar{K}$).

The proof and the definition of $\bar{\bar{K}}$ are in the appendix. q_H^* is the first-best level of output defined by $V'(q_H^*) = \theta_H + \tilde{\beta}$. We first interpret the proposition and then offer some intuition. The first part of the proposition considers the case where the gap between the outside opportunities of high and low skill agents is not too large ($\Delta U < \overline{\Delta U}$) and identifies the condition under which general training is more profitable for any given level of market frictions $\gamma \in (0,1)$. It shows that general training is more profitable for the principal than specific training for small improvements in the (after training) value of the agent's outside opportunities from general training ($K < \bar{K}$). Notice that \bar{K} is a fraction of ΔU ($\bar{K} \equiv \pi_H \Delta U$), implying that general training is more profitable than specific training for small K relative to ΔU ($K < \Delta U < \overline{\Delta U}$).

To understand the intuition behind the proposition, suppose first that $K = 0$. General training allows each type of the agent to identify the value of his outside opportunities, but does not increase it. If ΔU is small, the agent with high general skills still has an incentive to pretend to be a low general skill agent as in the case of specific training. However, since a low skill agent's outside opportunities are valued at a lower level after general training, pretending to be a low skill agent is not as attractive as in the case of specific training. It implies that disciplining the high skill agent's incentive becomes easier for the principal. In particular, the principal can reduce the information rent given to the high skill agent by $\gamma \Delta U$ compared to the case of specific training, making general training more profitable than specific training. However, if $K > 0$, the principal has to increase compensation for both types of the agent by γK . Then, general training is more profitable only if $\gamma K < \pi_H \gamma \Delta U$, which is the reduction of the expected information rent. Of course, this result is valid for small ΔU . If ΔU is large so that the agent with low general skills develops an incentive to pretend to be an agent with high general skills, the principal faces a new incentive problem, making general training less profitable than specific training.

The second part of the proposition shows that general training *can* still be more profitable than specific training for K larger than \bar{K} (i.e., $\bar{\bar{K}} > \bar{K}$). If ΔU is small, in addition to the reduction of the information rent given to the agent with high general skills, there is another gain from general training relative to specific training. Recall that the principal uses output distortion as well as the provision of rents when disciplining the agent's traditional incentive. With general training, the principal can ease output distortion since disciplining the high skill agent's incentive becomes easier. Depending on the magnitude of this additional benefit, general training can be more profitable for K larger than \bar{K} .

Notice that two necessary conditions for general training to be more profitable than specific training ($\Delta U < \overline{\Delta U}$ and $K < \bar{\bar{K}}$) depend on labor market frictions. In particular, greater frictions (smaller γ) make $\overline{\Delta U}$ and $\bar{\bar{K}}$ larger (see the definition of $\bar{\bar{K}}$ in the appendix), relaxing the conditions. Labor market frictions help general training to

be more profitable since the principal makes up only for a fraction of K , which decreases with the degree of labor market frictions γ , after general training.

Finally, other than the fact that an agent with high (low) general skills has the high (low) value of outside opportunities, we did not impose a specific relationship between the level of general skills and the value of outside opportunities. In particular, ΔU can have any values independent of those of $\Delta\beta$. A special case would be when the difference in the outside opportunities between high and low skill agents is closely related to rent inside the firm. It implies that $\Delta U \approx \Delta\beta \cdot q$. If this is the case, the condition $\Delta U < \overline{\Delta U} \equiv \frac{1}{\gamma} \Delta\beta q_H^*$ is likely to be met because $\gamma < 1$, and therefore the real binding condition depends on K only.

To summarize: because general training creates a force that counteracts the agent's traditional incentive of under-reporting his type, the principal can lessen the incentive problems and reduce the agent's rent. However, general training is costly to the principal since she has to compensate the agent for an increase in the value of his outside opportunities from general training. If the principal's gain from lessening incentive problems is larger than this cost, general training can be more profitable than specific training.

4. Extensions

In this section, we examine the robustness of our results to the cases where general and specific skills are positively correlated with each other, where the agent pays training costs, and where the (inside and outside) value of training is type-contingent. We also offer another application of our model in a different context.

4.1. Positive correlation between general and specific skills

When general and specific skills are correlated, even investment in specific training may reveal information about the agent's outside opportunities. A question then arises as to whether the result of over-investment in general training relative to specific training is robust to the introduction of a correlation between general and specific skills. If the correlation is perfect, there would no longer be any over-investment in general training relative to specific training because specific training becomes de facto general training. We show in this section that over-investment in general training relative to specific training remains as long as the correlation is not perfect.

To model the correlation between general and specific skills, we define conditional probabilities as $\pi_{ii} \equiv \text{prob}(\beta_i^S | \beta_i^G) = \text{prob}(\beta_i^G | \beta_i^S) = \delta + (1-\delta) \pi_i$ and $\pi_{ji} \equiv \text{prob}(\beta_j^S | \beta_i^G) = \text{prob}(\beta_j^G | \beta_i^S) = (1-\delta) \pi_j$, where $\delta \in (0,1)$ and $i \neq j \in \{H, L\}$.²² δ can represent the degree at which general and specific skills are correlated with each other.²³

²² Formally, the correlation between general and specific skills can be derived from joint probabilities of general and specific skills. However, what is required in our analysis is conditional probabilities in that an agent is more likely to have high specific skills given that he is an agent with high general skills, and vice versa.

²³ δ being zero implies that conditional probabilities become equivalent to unconditional probabilities ($\pi_{ii} = \pi_i$ and $\pi_{ji} = \pi_j$). π_{ii} increases whereas π_{ji} decreases as δ increases from zero, and finally they become 1 and 0 when $\delta = 1$, respectively.

An imperfect but positive correlation between general and specific skills implies that $\pi_{ii} > \pi_i$ and $\pi_{ji} < \pi_j$ for all $i \neq j$. With this positive correlation, the following proposition characterizes the condition under which over-investment in general training relative to specific training occurs.

Proposition 2 For $\delta \in (0,1)$, $\gamma \in (0,1)$, and $\Delta U < \overline{\Delta U}^\delta$, the principal always over-invests in general training relative to specific training if $K < \bar{K}^\delta$, where $\overline{\Delta U}^\delta = \frac{1}{\gamma} \Delta \beta q_L^1$ and $\bar{K}^\delta = (1-\delta)\pi_H \Delta U$.

The proof is in the appendix. q_L^1 is defined by $V'(q_L^1) = \theta_L + \tilde{\beta} + \frac{\pi_H}{\pi_L} \Delta \beta$. The proposition shows that regardless of labor market frictions and correlations between general and specific trainings, general training is more profitable for the principal than specific training for small K relative to ΔU . Labor market frictions help the principal to over-invest in general training relative to specific training as they ease the condition for general training to be more profitable ($\Delta U < \overline{\Delta U}^\delta$). An increase in the correlation δ reduces \bar{K}^δ , implying that general training becomes relatively less profitable with a higher correlation between general and specific trainings. In short, although general training becomes less attractive to the principal if there is a correlation between general and specific skills, it is still more profitable than specific training for small K relative to ΔU . The result of over-investment in general training relative to specific training is therefore robust to the introduction of the correlation.

We can offer the intuition behind the results as follows. Given that ΔU is small, both general and specific trainings lessen the agent's incentive problems. In particular, the high (general or specific) skill agent's incentive to pretend to be a low skill agent is weakened since by doing so he would be compensated based on a lower value of outside opportunities. The principal can thus reduce the information rent given to the high skill agent. As is well-known, the information rent given to the high skill agent is proportional to the compensation for the low skill agent and thereby to the value of the low skill agent's outside opportunities, which is $\gamma U_L + (1-\gamma)\tilde{U} + \gamma K$ in case of general training, $\gamma[\delta U_i + (1-\delta)\tilde{U}] + (1-\gamma)\tilde{U} = \delta\gamma U_L + (1-\delta\gamma)\tilde{U}$ in case of specific training, and \tilde{U} in case of no training. Both general and specific trainings reduce the information rent, and general training reduces it more than specific training does, making general training more profitable. However, the principal pays an extra cost of γK when investing in general training. This trade-off implies that general training is more profitable than specific training if $K < (1-\delta)\pi_H \Delta U$. Of course, a higher correlation between general and specific skills lowers the relative benefit of general training associated with rent reduction.

4.2. Training costs paid by agents

We have assumed so far that the principal pays training costs up front and that the agent has to participate in the training program regardless of his willingness to do so. In this section, we explore an alternative possibility where the agent pays training costs, and

only participates in the training program if it is profitable for him to do so.²⁴ Because our focus is not on the effects of the agent's credit constraints, we assume that the agent can access the fund for training costs without constraints.

Since the principal does not directly pay training costs, she must induce the agent to pay C and participate in the training program if it is profitable for her. Thus the principal has to face the following training participation constraint in addition to the contract participation constraint (IR) and the incentive compatibility constraint (IC) if training is directed toward specific skills:

$$\pi_H (t_H - (\tilde{\beta}^G + \theta_H^S)q_H) + \pi_L (t_L - (\tilde{\beta}^G + \theta_L^S)q_L) - \tilde{U} \geq C, \quad (3)$$

and if training is directed toward general skills:²⁵

$$\pi_H (t_H - (\theta_H^G + \tilde{\beta}^S)q_H) + \pi_L (t_L - (\theta_L^G + \tilde{\beta}^S)q_L) - \tilde{U} \geq C. \quad (4)$$

If the agent participates in the training program, he will receive his transfer according to his skills and therefore his expected payoff is the first two terms in the LHS of the constraints. If he does not participate in the program, his skills are not identified and therefore receives the payoff that is just equivalent to his expected outside opportunities \tilde{U} . To induce the agent to participate in the training program, the principal has to guarantee that the agent's gain from training (the LHS of the constraints) is larger than the cost of training C .

We can use the model of section 3 to examine how our solutions may be affected. Consider first specific training. Recall that the high skill agent is able to earn a rent due to the traditional incentive from private information. This makes the LHS of (3) positive. Thus, for small C , given that the principal offers the contract characterized in section 3, the agent voluntarily participates in the training program and pays C by himself, which is compensated by the rent he will later enjoy. However, for larger C , the training participation constraint would be violated. The principal has to pay some of the training costs by increasing the expected transfer to the agent ($\pi_L t_L + \pi_H t_H$).

Consider next general training. Depending on the size of ΔU , the agent may or may not earn a rent. However, the LHS of (4) is always positive since the principal has to compensate the agent for γK even if she does not have to surrender a rent. Thus, a similar result would be expected: for small C the agent pays all of the training costs, while for large C the principal pays some of the training costs. Since the effect of the training participation constraint is qualitatively the same for both specific and general trainings, the comparison of the level of investment in general training to specific training

²⁴ See WASMER [2006] that models the agent's incentive to invest in specific skills relative to general skills. In general, workers are more willing to invest in general skills than specific skills because of the exactly opposite reason why firms are less willing to invest in general skills. He investigates the conditions under which workers are more willing to invest in specific training. In a general equilibrium model with search frictions, he shows that workers invest more in specific skills relative to general skills when job matching is less efficient and when employment protection is high.

²⁵ In our model, the principal decides whether the training program is general or specific. Since our goal is to compare the relative profitability of the two types of training we do not consider a third option: investing in both general and specific training. Under such alternative assumption, the agent would need to determine not only his participation but also in which training program to join in, which would require additional constraints.

would be identical to the one in section 3.

To summarize: compared to the model in section 3, training is more attractive to the principal since she does not have to pay all of training costs. However, she has to pay some of the training costs if they become large. No matter who pays training costs, our main findings about the over-investment in general training relative to specific training remain.

4.3. Type contingent value of training

So far we have assumed that not only the “inside” value of training (α) but also the “outside” value of general training (K) are constant over the types. We relax these assumptions in this section and confirm the robustness of our results.

Let us first consider type-contingent values for α (i.e., $\alpha_H \neq \alpha_L$). Since training now affects the difference between the agent’s marginal costs, a new effect arises. Suppose that training benefits more the high skill agent: $\alpha_H > \alpha_L$. A direct consequence of training is to *increase* the gap between the marginal costs of the low and high skill agents. Training increases the informational rent of the high skill agent and less investment will occur, compared to a case where $\alpha_H = \alpha_L$. Conversely, if $\alpha_H < \alpha_L$, more investment will occur.

Notice that α was assumed to be the same for both general and specific skills for a comparison purpose ($\alpha^G = \alpha^S$), so it does not affect the relative profitability of general training to specific training no matter whether it is type-contingent or independent. However, the focus of this paper is on the incentive effects of a change in outside opportunities on the level of investment in training. Therefore, it is more relevant to examine how type contingent values for K (i.e., $K_H \neq K_L$) affect our results, putting aside type-contingent values for α .

We will start with the more natural case where $K_H > K_L$ (with $\Delta K \equiv K_H - K_L > 0$). This is the case where a high skill agent increases his outside opportunities relatively more with general training. For a comparison purpose, it is assumed that $K = \pi_H K_H + \pi_L K_L$. The difference between the high skill agent’s and the low skill agent’s outside opportunities is now $\Delta U + \Delta K$. The benefit to the principal of general training associated with rent reduction then depends on $\Delta U + \Delta K$ instead of ΔU alone, and the principal always over-invests in general training relative to specific training for $K < \pi_H(\Delta U + \Delta K)$. Compared to the case of constant K , general training is more profitable than specific training for greater values of K . The outside opportunities of the agent with low general skills were $\gamma U_L + (1-\gamma)\tilde{U} + \gamma K$, but they are now $\gamma U_L + (1-\gamma)\tilde{U} + \gamma K_L$ when K is type-contingent. Given that $K > K_L$, disciplining the high skill agent’s incentive to pretend to be a low skill agent becomes easier.

In contrast, if $K_H < K_L$ so that $\Delta K < 0$, general training becomes less attractive. However, our main result still remains. As long as the outside (after training) value of a high skill agent is greater than that of a low skill agent (i.e., $\Delta U + \Delta K > 0$), general training is more profitable than specific training for small K relative to $\Delta U + \Delta K$.

4.4. Investment in R&D

Our model and findings can also be applied to other areas such as investment in

R&D. Consider for instance a firm that invests in research to improve and develop projects. As a result of R&D investment, a researcher (the agent) acquires a new technology or project idea, which enhances efficiency inside the firm (α). A successful result of R&D investment is likely to be related with the researcher's ability. Thus, at the time R&D investment takes place, the researcher learns his ability (β). Just like investment in training, investment in R&D also has dual roles: it increases inside efficiency and identifies the researcher's ability.

The researcher with high ability may expect better outside opportunities ($U_H > U_L$) if the technology or the project idea the researcher developed is general rather than firm-specific. In addition, the researcher may have new outside opportunities (K). For instance, he could walk away with the technology or project idea and develop it elsewhere as argued in ANTON AND YAO [1995] and ANAND AND GALETOVIC [2000]. The size of new outside opportunities depends on not only the outside value of the result of R&D investment but also how well property right is defined on it.

Our model of investment in information gathering produces a conventional prediction regarding the effect of property right on investment in R&D.²⁶ If property rights on the result of research are weak (large K), the benefit from easing incentive problems through general R&D investment is not enough to offset the negative effect of the outflow of research results. It leads to under-investment in general R&D investment relative to firm specific R&D investment. If property rights are strong (small K), the benefit of lessening incentive problems through general R&D investment can dominate. Then, there may be over-investment in general R&D.

5. Conclusion

Our model studied an important effect of the identification of outside opportunities on investment in information gathering. This effect has unexpected consequences when information revelation is asymmetric: it generates over-investment in information gathering relative to the case where information gathering does not identify the agent's outside opportunities.

We presented an application of our model to investment in general training for new hires. We showed that an employer may over-invest in general training relative to specific training. Becker's traditional result of under-investment in general training may not hold when training creates asymmetric information. The identification of an agent's outside opportunities from general training lessens the incentive problems inside the firm. This is a novel justification of the evidence reported in the introduction that employers sponsor general training particularly for young workers, new hires and apprentices. Our explanation relies on incentives *inside* the firm. Except for the paper by KESSLER AND LÜLFESMANN [2006] so far the literature had only stressed causes *outside* the firm such as labor market frictions.

Our model offers a few interesting policy implications. As is well-known, the presence of information asymmetry in economic trades results in sub-optimal outcome and thereby is welfare-decreasing. Thus policies that reduce the information gap such as monitoring, auditing, certification, and others are recommended in many instances. However, in the context of general training, our model shows that the elimination of

²⁶ See AGHION AND TIROLE [1994] for optimal allocation of property rights on the result of research.

information asymmetry may make general training less likely. In the case where information asymmetry disappears because of the firm's learning by doing or other natural reasons (for example the case of senior workers), training subsidies would help employers to sponsor general training.

Appendix

Investment in training under symmetric information

With specific training that produces symmetric information, the contract is based on θ_i^S , $\tilde{\beta}^G$, and \tilde{U} . The principal's problem becomes

$$\begin{aligned} & \text{Max}_{q_i, t_i} \sum_{i \in \{H, L\}} \pi_i [V(q_i) - t_i] - C \\ & \text{s.t. } (IR_i) \quad t_i - (\tilde{\beta}^G + \theta_i^S)q_i \geq \tilde{U} \quad \forall i \in \{H, L\}. \end{aligned}$$

The optimal contract requires the ex-post first-best level of output, denoted by q_i^{S*} , for each type of skill: $V'(q_i^{S*}) = \tilde{\beta}^G + \theta_i^S$, $i = H, L$. The transfer given to the agent is $t_i = (\tilde{\beta}^G + \theta_i^S)q_i^{S*} + \tilde{U}$, $i = H, L$. Therefore, at the optimum, the value of the objective function gross of C is

$$W^S = \sum_{i \in \{H, L\}} \pi_i [V(q_i^{S*}) - (\tilde{\beta}^G + \theta_i^S)q_i^{S*}] - \tilde{U} \quad (\text{A1})$$

Investment in specific training will take place if its cost, C , is less than its benefit $C^S = W^S - W^{NT}$:

$$C^S = \sum_{i \in \{H, L\}} \pi_i [V(q_i^{S*}) - (\tilde{\beta}^G + \theta_i^S)q_i^{S*}] - [V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S)\tilde{q}] \quad (\text{A2})$$

Clearly, the principal can be better off by investing in specific training ($C^S > 0$). This is due to both the information gathering and the human capital functions of training.

Next, when the principal invests in general training, the contract is now based on θ_i^G , $\tilde{\beta}^S$, and $\gamma U_i + (1 - \gamma)\tilde{U} + \gamma K$. The principal's problem becomes

$$\begin{aligned} & \text{Max}_{q_i, t_i} \sum_{i \in \{H, L\}} \pi_i [V(q_i) - t_i] - C \\ & \text{s.t. } (IR_i) \quad t_i - (\theta_i^G + \tilde{\beta}^S)q_i \geq \gamma U_i + (1 - \gamma)\tilde{U} + \gamma K \quad \forall i \in \{H, L\}. \end{aligned}$$

Both types of the agent still produce at the first best: $V'(q_i^{G*}) = \theta_i^G + \tilde{\beta}^S$, $i = H, L$. But investment in training is affected as it now depends on γK . Labeling the optimal value of the objective function gross of C as W^G , the principal invests in training if C is less than $C^G = W^G - W^{NT}$:

$$C^G = \sum_{i \in \{H, L\}} \pi_i [V(q_i^{G*}) - (\theta_i^G + \tilde{\beta}^S)q_i^{G*}] - [V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S)\tilde{q}] - \gamma K \quad (\text{A3})$$

If $\beta_i^G = \beta_i^S \equiv \beta_i$ and $\alpha^G = \alpha^S \equiv \alpha$, it must be that $\theta_i^G = \theta_i^S \equiv \theta_i$ and $\tilde{\beta}^G = \tilde{\beta}^S \equiv \tilde{\beta}$, implying that $q_i^{S*} = q_i^{G*} \equiv q_i^*$. From (A2) and (A3), it is then clear that the principal under-invests in general training relative to specific training ($C^G < C^S$).

Proof of proposition 1

LAFFONT AND TIROLE [1989] and LAFFONT AND MARTIMORT [2002] show that in the case of type-contingent outside opportunities, the optimal contract has five different regimes depending on the size of ΔU . We first summarize the optimal outputs of each regime.

Regime	Region	Optimal output
I	$\Delta U < \frac{1}{\gamma} \Delta \beta^G q_L^{G1}$	$V'(q_L^{G1}) = \theta_L^G + \tilde{\beta}^S + \frac{\pi_H}{\pi_L} \Delta \beta^G$, $V'(q_H^{G1}) = \theta_H^G + \tilde{\beta}^S$
II	$\frac{1}{\gamma} \Delta \beta^G q_L^{G1} < \Delta U < \frac{1}{\gamma} \Delta \beta^G q_L^{G3}$	$V'(q_L^{G2}) = \theta_L^G + \tilde{\beta}^S + \frac{\lambda_1}{\pi_L} \Delta \beta^G$, $V'(q_H^{G2}) = \theta_H^G + \tilde{\beta}^S$
III	$\frac{1}{\gamma} \Delta \beta^G q_L^{G3} < \Delta U < \frac{1}{\gamma} \Delta \beta^G q_H^{G3}$	$V'(q_L^{G3}) = \theta_L^G + \tilde{\beta}^S$, $V'(q_H^{G3}) = \theta_H^G + \tilde{\beta}^S$
IV	$\frac{1}{\gamma} \Delta \beta^G q_H^{G3} < \Delta U < \frac{1}{\gamma} \Delta \beta^G q_H^{G5}$	$V'(q_L^{G4}) = \theta_L^G + \tilde{\beta}^S$, $V'(q_H^{G4}) = \theta_H^G + \tilde{\beta}^S - \frac{\lambda_2}{\pi_H} \Delta \beta^G$
V	$\Delta U > \frac{1}{\gamma} \Delta \beta^G q_H^{G5}$	$V'(q_L^{G5}) = \theta_L^G + \tilde{\beta}^S$, $V'(q_H^{G5}) = \theta_H^G + \tilde{\beta}^S - \frac{\pi_L}{\pi_H} \Delta \beta^G$

Here $\lambda_1 \leq \pi_H$ is the multiplier for the (IC_H) constraint, and $\lambda_2 \leq \pi_L$ is the multiplier for the (IC_L) constraint. Thus,

$$\begin{aligned} q_L^{G1} &\leq q_L^{G2} \leq q_L^{G3} = q_L^{G4} = q_L^{G5} \equiv q_L^{G*} \\ q_H^{G*} &\equiv q_H^{G1} = q_H^{G2} = q_H^{G3} \leq q_H^{G4} \leq q_H^{G5}. \end{aligned}$$

The principal's maximum willingness to pay for general training, C_r^G , $r = 1, \dots, 5$, is $W_r^G - W^{NT}$, where W_r^G is the value of the objective function gross of C at regime r when investment in general training takes place while W^{NT} is the value of the objective function when no investment takes place. Thus, C_r^G s are given by

$$\begin{aligned} C_1^G &= \pi_H [V(q_H^{G1}) - (\theta_H^G + \tilde{\beta}^S) q_H^{G1}] + \pi_L [V(q_L^{G1}) - (\theta_L^G + \tilde{\beta}^S) q_L^{G1}] \\ &\quad - [V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S) \tilde{q}] - \pi_H \Delta \beta^G q_L^{G1} + \pi_H \gamma \Delta U - \gamma K \\ C_2^G &= \pi_H [V(q_H^{G2}) - (\theta_H^G + \tilde{\beta}^S) q_H^{G2}] + \pi_L [V(q_L^{G2}) - (\theta_L^G + \tilde{\beta}^S) q_L^{G2}] \\ &\quad - [V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S) \tilde{q}] - \gamma K \\ C_3^G &= \pi_H [V(q_H^{G3}) - (\theta_H^G + \tilde{\beta}^S) q_H^{G3}] + \pi_L [V(q_L^{G3}) - (\theta_L^G + \tilde{\beta}^S) q_L^{G3}] \\ &\quad - [V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S) \tilde{q}] - \gamma K \end{aligned}$$

$$\begin{aligned}
C_4^G &= \pi_H \left[V(q_H^{G4}) - (\theta_H^G + \tilde{\beta}^S) q_H^{G4} \right] + \pi_L \left[V(q_L^{G4}) - (\theta_L^G + \tilde{\beta}^S) q_L^{G4} \right] \\
&\quad - \left[V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S) \tilde{q} \right] - \gamma K \\
C_5^G &= \pi_H \left[V(q_H^{G5}) - (\theta_H^G + \tilde{\beta}^S) q_H^{G5} \right] + \pi_L \left[V(q_L^{G5}) - (\theta_L^G + \tilde{\beta}^S) q_L^{G5} \right] \\
&\quad - \left[V(\tilde{q}) - (\tilde{\beta}^G + \tilde{\beta}^S) \tilde{q} \right] + \pi_L \Delta \beta q_H^{G5} - \pi_L \gamma \Delta U - \gamma K
\end{aligned}$$

With symmetric conditions, $\beta_i^G = \beta_i^S \equiv \beta_i$ and $\alpha^G = \alpha^S \equiv \alpha$, it must be that $\theta_i^G = \theta_i^S \equiv \theta_i$, $\tilde{\beta}^G = \tilde{\beta}^S \equiv \tilde{\beta}$, and $\Delta \beta^G = \Delta \beta^S \equiv \Delta \beta$. It is then easy to see that $q_L^S = q_L^{G1} \leq q_L^{G*} \equiv q_L^*$ and $q_H^S = q_H^{G*} \equiv q_H^*$.

Next, we define the following useful expression.

$$Y(q_i) \equiv V(q_i) - (\theta_i + \tilde{\beta}) q_i, \quad i = H, L.$$

Since $V(\cdot)$ is concave and $(\theta_i + \tilde{\beta}) q_i$ is linear in q_i , $Y(q_i)$ has an unique maximum at $q_i = q_i^*$, and it is monotonically decreasing as q_i goes away from q_i^* . It implies that

$$Y(q_L^*) \geq Y(q_L^{G2}) \geq Y(q_L^{G1}) = Y(q_L^S), \quad Y(q_H^*) \geq Y(q_H^{G4}) \geq Y(q_H^{G5}), \quad (\text{A4})$$

since $q_L^* \geq q_L^{G2} \geq q_L^{G1} = q_L^S$ and $q_H^* \leq q_H^{G4} \leq q_H^{G5}$.

Now we are ready to compare C_r^G , $r=1, \dots, 5$, with C^S .

$$\begin{aligned}
C_1^G - C^S &= \pi_H \gamma \Delta U - \gamma K \\
C_2^G - C^S &= \pi_L \left[Y(q_L^{G2}) - Y(q_L^S) \right] + \pi_H \Delta \beta q_L^S - \gamma K \\
C_3^G - C^S &= \pi_L \left[Y(q_L^*) - Y(q_L^S) \right] + \pi_H \Delta \beta q_L^S - \gamma K \\
C_4^G - C^S &= -\pi_H \left[Y(q_H^*) - Y(q_H^{G4}) \right] + \pi_L \left[Y(q_L^*) - Y(q_L^S) \right] + \pi_H \Delta \beta q_L^S - \gamma K \\
C_5^G - C^S &= -\pi_H \left[Y(q_H^*) - Y(q_H^{G5}) \right] + \pi_L \left[Y(q_L^*) - Y(q_L^S) \right] + \pi_H \Delta \beta q_L^S + \pi_L \Delta \beta q_H^{G5} - \pi_L \gamma \Delta U - \gamma K
\end{aligned}$$

where all of the expressions in [] are non-negative from (A4). Notice that if $K = 0$, $C_1^G > C^S$, $C_2^G > C^S$, and $C_3^G > C^S$. However, C_4^G and C_5^G may or may not be greater than C^S . Thus, we can find sufficient conditions associated with K that allow C^G to be greater than C^S if $\Delta U < \overline{\Delta U} \equiv \frac{1}{\gamma} \Delta \beta q_H^*$ (from regime 1 to regime 3). In particular,

$$\begin{aligned}
C_1^G &> C^S \text{ if } K < \bar{K} \equiv \pi_H \Delta U \\
C_2^G &> C^S \text{ if } K < \hat{K} \equiv \frac{1}{\gamma} \left\{ \pi_L \left[Y(q_L^{G2}) - Y(q_L^S) \right] + \pi_H \Delta \beta q_L^S \right\} \\
C_3^G &> C^S \text{ if } K < \bar{\bar{K}} \equiv \frac{1}{\gamma} \left\{ \pi_L \left[Y(q_L^*) - Y(q_L^S) \right] + \pi_H \Delta \beta q_L^S \right\}.
\end{aligned}$$

Since $\Delta U < \frac{1}{\gamma} \Delta \beta q_L^S$ in regime 1, $\bar{K} < \hat{K} < \bar{\bar{K}}$. Thus, given that $\Delta U < \bar{\Delta U}$, C^G is always greater than C^S if $K < \bar{K}$, and C^S can be greater than C^S if $K < \bar{\bar{K}}$. ■

Proof of proposition 2

Given that there is a positive correlation between general and specific skills, the principal faces the following problem when investing in specific training:

$$\begin{aligned} & \text{Max}_{q_i, t_i} \sum_{i \in \{H, L\}} \pi_i [V(q_i) - t_i] - C \\ \text{s.t. } & (IR_i) \quad t_i - (\tilde{\beta}_i^G + \theta_i^S) q_i \geq \delta \gamma U_i + (1 - \delta \gamma) \tilde{U} \quad \forall i \in \{H, L\} \\ & (IC_i) \quad t_i - (\tilde{\beta}_i^G + \theta_i^S) q_i \geq t_j - (\tilde{\beta}_j^G + \theta_j^S) q_j \quad \forall i, j \in \{H, L\} \end{aligned}$$

Similar to the one shown in the proof of proposition 1, the solution to this problem is summarized as:

Regime	Region	Optimal output
I	$\Delta U < \frac{1}{\delta \gamma} \Delta \beta^S q_L^{S1}$	$V'(q_L^{S1}) = \tilde{\beta}_L^G + \theta_L^S + \frac{\pi_H}{\pi_L} \Delta \beta^S$, $V'(q_H^{S1}) = \tilde{\beta}_H^G + \theta_H^S$
II	$\frac{1}{\delta \gamma} \Delta \beta^S q_L^{S1} < \Delta U < \frac{1}{\delta \gamma} \Delta \beta^S q_L^{S3}$	$V'(q_L^{S2}) = \tilde{\beta}_L^G + \theta_L^S + \frac{\lambda_1}{\pi_L} \Delta \beta^S$, $V'(q_H^{S2}) = \tilde{\beta}_H^G + \theta_H^S$
III	$\frac{1}{\delta \gamma} \Delta \beta^S q_L^{S3} < \Delta U < \frac{1}{\delta \gamma} \Delta \beta^S q_H^{S3}$	$V'(q_L^{S3}) = \tilde{\beta}_L^G + \theta_L^S$, $V'(q_H^{S3}) = \tilde{\beta}_H^G + \theta_H^S$
IV	$\frac{1}{\delta \gamma} \Delta \beta^S q_H^{S3} < \Delta U < \frac{1}{\delta \gamma} \Delta \beta^S q_H^{S5}$	$V'(q_L^{S4}) = \tilde{\beta}_L^G + \theta_L^S$, $V'(q_H^{S4}) = \tilde{\beta}_H^G + \theta_H^S - \frac{\lambda_2}{\pi_H} \Delta \beta^S$
V	$\Delta U > \frac{1}{\delta \gamma} \Delta \beta^S q_H^{S5}$	$V'(q_L^{S5}) = \tilde{\beta}_L^G + \theta_L^S$, $V'(q_H^{S5}) = \tilde{\beta}_H^G + \theta_H^S - \frac{\pi_L}{\pi_H} \Delta \beta^S$

Here $\lambda_1 \leq \pi_H$ is the multiplier for the (IC_H) constraint, and $\lambda_2 \leq \pi_L$ is the multiplier for the (IC_L) constraint.

When investing general training, the principal's problem can be stated as

$$\begin{aligned} & \text{Max}_{q_i, t_i} \sum_{i \in \{H, L\}} \pi_i [V(q_i) - t_i] - C \\ \text{s.t. } & (IR_i) \quad t_i - (\theta_i^G + \tilde{\beta}_i^S) q_i \geq \gamma U_i + (1 - \gamma) \tilde{U} + \gamma K \quad \forall i \in \{H, L\} \\ & (IC_i) \quad t_i - (\theta_i^G + \tilde{\beta}_i^S) q_i \geq t_j - (\theta_j^G + \tilde{\beta}_j^S) q_j \quad \forall i, j \in \{H, L\}. \end{aligned}$$

Notice that this problem and its solution are virtually the same as the ones in the case where $\delta = 0$, except that specific skills are now $\tilde{\beta}_i^S$ instead of $\tilde{\beta}_i^S$.

With symmetric conditions, $\beta_i^G = \beta_i^S \equiv \beta_i$ and $\alpha^G = \alpha^S \equiv \alpha$, it must be that

$\theta_i^G = \theta_i^S \equiv \theta_i$, $\tilde{\beta}^G = \tilde{\beta}^S \equiv \tilde{\beta}$, and $\Delta\beta^G = \Delta\beta^S \equiv \Delta\beta$. It is then easy to see that $q_i^{Sr} = q_i^{Gr} \equiv q_i^r$, $i = H, L$, and $r = 1, \dots, 5$. Using these solutions, we can calculate the principal's maximum willingness to pay for specific and general trainings, C_r^S and C_r^G , $r = 1, \dots, 5$, as shown in the proof of proposition 1. Notice that the region of each regime does not match between the cases of specific and general training. However, for $\Delta U < \frac{1}{\gamma} \Delta\beta q_L^1$, the optimal contract in both specific and general training cases are in regime 1, and

$$C_1^G - C_1^S = (1 - \delta)\pi_H \gamma \Delta U - \gamma K.$$

Given that $\Delta U < \overline{\Delta U}^\delta \equiv \frac{1}{\gamma} \Delta\beta q_L^1$, the above comparisons imply that $C_1^G > C_1^S$ if $K < \overline{K}^\delta \equiv (1 - \delta)\pi_H \Delta U$. ■

References

- ACEMOGLU, D. [1997], "Training and Innovation in an Imperfect Labour Market," *Review of Economic Studies*, 64, 445-464.
- – AND J. PISCHKE [1998], "Why Do Firms Train? Theory and Evidence," *Quarterly Journal of Economics*, 113, 79-119.
- – AND – – [1999], "The Structure of Wage and Investment in General Training," *Journal of Political Economy*, 107, 539-572.
- – AND – – [2000], "Certification of Training and Training Outcomes," *European Economic Review*, 44, 917-927.
- AGHION, P. AND J. TIROLE [1994], "The Management of Innovation," *Quarterly Journal of Economics*, 109, 1185-1209.
- ANAND, B. AND A. GALETOVIC [2000], "Weak Property Right and Holdup in R&D," *Journal of Economics and Management Strategy*, 9, 615-642.
- ANTON, J. AND D. YAO [1995], "Start-ups, Spin-offs, and Internal Projects," *Journal of Law, Economics and Organization*, 11, 362-378.
- AUTOR, D. [2001], "Why Do Temporary Help Firms Provide Free General Skills Training?" *Quarterly Journal of Economics*, 116, 1409-1448.
- BECKER, G. [1964], *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, Columbia University Press: New York, NY.
- BONTEMS, P. AND J. BOURGEON [2000], "Creating Countervailing Incentives through the Choice of Instruments," *Journal of Public Economics*, 76, 181-202.
- BOOTH, A. AND M. BRYAN [2007], "Who Pays General Training in Private Sector Britain?" *Research in Labor Economics*, 28, 83-121.
- BUDDIN, R. AND K. KAPUR [2005], "The Effect of Employer-Sponsored Education on Job Mobility," *Industrial Relations*, 44, 341-363.
- CHANG, C. AND Y. WANG [1996], "Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited," *Journal of Labor Economics*, 14, 505 – 519.

- CRÉMER, J. AND F. KHALIL [1992], "Gathering Information before Signing a Contract," *American Economic Review*, 82, 566-578.
- –, – –, AND J. ROCHET [1998a], "Contracts and Productive Information Gathering," *Games and Economic Behavior*, 25, 174-193.
- –, – –, AND – – [1998b], "Strategic Information Gathering Before a Contract is Offered," *Journal of Economic Theory*, 81, 163-200.
- DEWATRIPONT, M., I. JEWITT, AND J. TIROLE [1999], "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, 66, 183-198.
- GEHRIG, T. [2004], "Information Acquisition and Organisational Form," *Journal of Institutional and Theoretical Economics*, 160, 1-13.
- GERFIN, M. [2004], "Firm-Sponsored General Training in Frictional Labour Markets: An Empirical Analysis for Switzerland," IZA Discussion paper, No. 1077, Institute for the Study of Labor, Bonn.
- HARHOFF, D. AND T. KANE [1997], "Is the German Apprenticeship System a Panacea for the U.S. Labor Market?" *Journal of Population Economic*, 10, 171-196.
- HECKMAN, J., R. ROSELIUS, AND J. SMITH [1994], "U.S. Education and Training Policy: A Re-evaluation of the Underlying Assumptions behind the New Consensus," pp.83-121 in L. Solomon and A. Levenson (eds.), *Labor markets, employment policy, and job creation*, Westview Press, Boulder, CO.
- HOLMSTRÖM, B. [1982], "Managerial Incentive Problems: A Dynamic Perspective," in *Essays in Economics and Management in Honor of Lars Wahlbeck*, Swedish School of Economics, Helsinki; Reprinted in *Review of Economic Studies*, 66, 169-182.
- JULLIEN, B. [2000], "Participation Constraints in Adverse Selection Models," *Journal of Economic Theory*, 93, 1-47.
- KATZ, E. AND A. ZIDERMAN [1990], "Investment in General Training: The Role of Information and Labour Mobility," *Economic Journal*, 100, 1147-1158.
- KESSLER, A. [1998], "The Value of Ignorance," *Rand Journal of Economics* 29, 339-354.
- – AND C. LÜLFESMANN [2006], "The Theory of Human Capital Revisited: On the Interaction of General and Specific Investments," *Economic Journal*, 116, 903-923.
- LAFFONT, J. AND D. MARTIMORT [2002], *The Theory of Incentives*, Princeton University Press: Princeton, NJ.
- – AND J. TIROLE [1989], "Optimal Bypass and Cream Skimming," *American Economic Review*, 80, 1042-1061.
- – AND – – [1993], *A Theory of Incentives in Procurement and Regulation*, MIT Press: Cambridge, MA.
- LAZEAR, E. AND S. ROSEN [1981], "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841-864.
- LEWIS, T. AND D. SAPPINGTON [1989a], "Countervailing Incentives in Agency Problems," *Journal of Economic Theory*, 49, 294-313.
- – AND – – [1989b], "Inflexible Rules in Incentive Problems," *American Economic Review*, 79, 69-84.
- – AND – – [1993], "Choosing Workers' Qualifications: No Experience Necessary," *International Economic Review*, 34, 479-502.

- – AND – – [1994], “Supplying Information to Facilitate Price Discrimination,” *International Economic Review*, 35, 309-327.
- – AND – – [1997], “Information Management in Incentive Problems,” *Journal of Political Economy*, 105, 796-821.
- LOWENSTEIN, M. AND J. SPLETZER [1998], “Dividing the Costs and Returns to General Training,” *Journal of Labor Economics*, 16, 142-171.
- LYNCH, L. [1992], “Private-Sector Training and the Earnings of Young Workers,” *American Economic Review*, 82, 299-312.
- MASKIN, E. AND J. TIROLE [1990], “The Principal-Agent Relationship with an Informed Principal, I: Private Values,” *Econometrica*, 58, 379-410.
- – AND – – [1992], “The Principal-Agent Relationship with an Informed Principal, II: Common Values,” *Econometrica*, 60, 1-42.
- STEVENS, M. [1994], “Labour Contracts and Efficiency in on-the-Job Training,” *Economic Journal*, 104, 408-419.
- U.S. BUREAU OF LABOR STATISTICS [2007], “National Compensation Survey: Employee Benefits in Private Industry in the United States, March 2007,” <http://www.bls.gov/ebs/>.
- WASMER, E. [2006], “General versus Specific Skills in Labor Markets with Search Frictions and Firing Costs,” *American Economic Review*, 96, 810-831.