Firm-Specific Training

LEONARDO FELLI (London School of Economics)

CHRISTOPHER HARRIS (King's College, Cambridge)

First Draft: March 12, 2004*
This Draft: March 28, 2004

ABSTRACT. This paper introduces two complementary models of firm-specific training: an informational model and a productivity-enhancement model. In both models, market provision of firm-specific training is inefficient. However, the nature of the inefficiency depends on the balance between the two key components of training, namely productivity enhancement and employee evaluation. In the informational model, training results in a proportionate increase in productivity enhancement and employee evaluation, and training is underprovided by the market. In the productivity-enhancement model, training results in an increase in productivity enhancement but no change in employee evaluation, and training is overprovided by the market. In both models, turnover is inefficiently low.

1. Introduction

A widely documented feature of the US labour market is the high mobility of young workers. A typical male worker will hold seven jobs during his first ten years in the labour market. This number amounts to about two thirds of the total number of jobs he holds during his entire career.[1]

There is a positive aspect to this mobility: job-shopping early in a worker's career may help him to settle into a good match relatively quickly. The worker does not therefore spend too much time accumulating human capital specific to a bad match. There is also a negative aspect: job-shopping early in a worker's career may enable him to accumulate a small amount of firm-specific human capital in each of a large number of firms, but prevent him from accumulating a significant amount of firm-specific human capital in any one firm.[2]

In a world in which the accumulation of firm-specific human capital is passive, it can be argued that the market will achieve the optimal trade off between these two aspects of mobility. Indeed, in such a world, the principal decision is made by the

*We are grateful to Steve Nickell for encouraging us to think about training, to Heski Bar-Isaac, Chris Pissarides, Steve Pischke, Yona Rubinstein and Yoram Weiss for very useful discussions and to Malte Loos for research assistance. Leonardo Felli would like to thank the Department of Economics of the Stern School of Business at NYU for their generous hospitality. Christopher Harris would like to thank the British Academy for their generous support, and the Institute for Advanced Studies, Princeton for their generous hospitality.

[1]See Topel and Ward (1992).

[2]Cf. Section 4B of Heckman (1993).

worker, who must choose his employer. Moreover, firms can influence this choice via their wage offers. The mobility decisions of the worker should therefore be socially efficient.[3]

In a world in which the accumulation of firm-specific human capital is active, the situation is more complicated. There are now two decisions to be made: the worker must choose his employer; and the employer must choose whether or not to train the worker. Moreover, while firms can still influence the worker's choice of employer via their wage offers, a non-employer cannot influence the training choices of an employer. Training choices cannot therefore be expected to be socially efficient.[4]

The purpose of the present paper is to investigate the market provision of firm-specific training, and to analyze the inefficiencies associated with it. To this end, we introduce two complementary models, namely an informational model and a productivity-enhancement model.

In both models, there are two firms and one worker. In each period, the two firms compete for the worker on the basis of her expected productivity in the two competing matches. Each firm offers her a two-part contract, which specifies (i) her wage and (ii) whether she will be assigned to the job or to a training program. The worker then chooses between the two offers, and undertakes the assignment specified in her contract. This results in a change in the expected productivity of the match. We think of the mean and variance of the change as the *productivity-enhancement* and *employee-evaluation* components of learning-by-doing respectively. Where the two models differ is in the impact that training has on these two components of learning-by-doing.[5]

---

[3]See Felli and Harris (1996). See also Bergemann and Välimäki (1996) for a closely related model that looks at a different application, namely sellers competing for a buyer through dynamic pricing.

[4]Prendergast (1993) develops a model where an employer can create efficient incentives for a worker to (actively) accumulate firm-specific human capital by committing to a pay scale that associates different remunerations to different tasks associated, in turn, with different levels of firm-specific human capital. The key difference between the analysis in Prendergast (1993) and our analysis is the ability of the employer to commit to a pay scale. In our environment employers cannot commit to a long term contract. Therefore, the only mechanism through which in equilibrium different levels of firms-specific human capital can be associated to different wage rates is the employers' competition for the worker.

[5]Case-study evidence seems to suggest that employee evaluation is a key component of the firm-specific-training programs offered by some European companies. For example, the Association of Retailers in France provides a rather extended period of training for new employees. The training is formal, and trainees that succeed in the final exam are awarded a professional diploma. Employees are initially selected by individual retailers (normally supermarkets) and then enrolled in the training program. After the training program employees sometimes end up changing retailers. In other words, training does seem to foster worker's mobility. (Cf. The Retail Sector in France: Report for the Force Programme, CEDEFOP, Berlin 1993.) Another example is the training provided by Mercedes Benz Car Dealers in Germany. The employees are offered a whole range of training courses by the employers. New employees are offered basic introductory training and highly technical training programs are offered to the specialized work force of the company through the Mercedez Benz training center. Training is clearly aimed at rendering the participants fully familiar with new car models and the human capital accumulated in these courses is highly specific since it dies when the

In the informational model, we model firm-specific human capital as information about the match between a firm and the worker; and we assume that, if the firm employs the worker for a period and assigns her to the job, then this will generate an increment in the information about the match.[6] It is then natural to go on to assume that, if the firm employs the worker and assigns her to the training program, then this will result in a larger increment in the information about the match. In other words, firm-specific human capital accumulates at a faster rate when the worker is assigned to the training program than it does when the worker is assigned to the job.

Incremental information about the match can be used in two ways. First, it can be used to improve the assignment of the worker within the firm. Such reassignment will, on average, raise the productivity of the worker. In other words, it will result in productivity enhancement. Secondly, the information can be used to re-evaluate the quality of the match between the worker and the firm. If the worker's productivity has gone up, then the value of the worker to the firm (and hence the willingness of the firm to pay for the worker) will be revised up. (This will ensure that the worker remains with the current employer.) If instead the worker's productivity has gone down, then the value of the worker to the firm (and hence the willingness of the firm to pay for the worker) will be revised down. (This might lead the worker to choose a different employer.) In other words, employee evaluation takes place.

It follows that, if the firm assigns the worker to the job, then this will result in both productivity enhancement and employee evaluation; and, if the firm switches the worker from the job to the training program, then this will result in an increase in both productivity enhancement and employee evaluation. Indeed, we show that the switch will result in a *proportionate* increase in the two components of learning-by-doing.

In the productivity-enhancement model, we do not attempt to build a microfoundation for productivity enhancement and employee evaluation. We simply assume that, if the firm switches the worker from the job to the training program, then productivity enhancement will increase but employee evaluation will remain unchanged.

In both models, both firms are free to adjust the wage element of their offers. The employer therefore internalizes the preferences of the worker as to whether she should be assigned to the job or to the training program. However, the other firm has no way to express its preference as to whether the employer should assign the worker to the job or to the training program. In order to determine whether training is underprovided or overprovided, we therefore need only determine whether the other firm assigns a positive or a negative value to training by the employer.

In the informational model, the key uncertainty at any point in time is whether the

model is taken out of production. Evidence suggests that mobility following the introductory course is particularly high: on average, only one in six trainees are retained as employees. (Cf. Motor Vehicle Repair and Sales Sector: Germany Report for the Force Programme, Berlin 1993.) This can be interpreted as evidence that employee evaluation is a relevant component of the training program.

[6] Cf. Jovanovic (1979) and Felli and Harris (1996).

worker will remain forever with her current employer, or whether she will eventually move to the other firm. Training accelerates the accumulation of information about the match with the current employer, and therefore brings forward the time at which this uncertainty is resolved. Specifically, if the current employer switches the worker from the job to the training program, then: in those states of the world in which the worker originally remained forever with her current employer, she will still remain forever with her current employer; and, in those states of the world in which she eventually moved to the other firm, she will move to the other firm sooner. In the first case, the other firm neither gains nor loses. In the second case, the other firm gains. Indeed, it can put the worker to productive use sooner. Training by the current employer is therefore unambiguously good from the point of view of the other firm, and training is underprovided in equilibrium. Moreover, given that the benefit of training to the other firm is precisely that it causes the worker to change employer sooner, turnover is inefficiently low.

In the productivity-enhancement model, training has the effect of raising the entire future timepath of the worker's productivity with the current employer. This has two effects. First, in some cases in which the worker was previously employed by the other firm, she will now be employed by the current employer. Secondly, even in cases where she is still employed by the other firm, the other firm will now have to pay her a higher wage, since her outside option will be higher. In the first case, the other firm loses an opportunity to put the worker to productive use. In the second case, the other firm has to pay her more. Training by the current employer is therefore unambiguously bad from the point of view of the other firm, and training is overprovided in equilibrium.[7] Moreover, given that part of the cost of training to the other firm is that it prevents the worker from switching employer in some cases, turnover is inefficiently low.[8]

The conclusions drawn from our two alternative models of training differ in terms of the nature of the inefficiency that arises in equilibrium: the informational model predicts that training is underprovided; whereas the productivity-enhancement model predicts that training is overprovided. However, the difference between these conclu-

[7] The inefficiency in the productivity-enhancement model is similar to that in Pissarides (1994). In that model, workers underinvest in on-the-job search because they do not take into account the positive externality that this search activity exerts on the other firm. As a result, turnover is inefficiently low and on-the-job accumulation of firm-specific skills is inefficiently high.

[8] The argument that firm-specific training may inhibit turnover is already present in Becker (1993). Becker suggests two reasons why this may be the case. First, the marginal product of a worker who possesses firm-specific human capital may exceed her wage. This implies that such a worker is more likely to be retained in the face of an exogenous downward shift in productivity (or an exogenous upward shift in wages) than a worker who possesses only general human capital. Secondly, even if the marginal product of such a worker falls below her wage, the firm may still choose to retain her. This is because, if the firm lets her go during a downturn, it may be unable to rehire her during a subsequent upturn. In other words, there is an option value associated with a worker who possesses firm-specific human capital (Becker 1993, Chapter III.1). Becker's is, however, a partial equilibrium model, and he does not therefore comment on whether turnover is efficient.

sions is attributable to the different balance between the productivity-enhancement and employee-evaluation components of training. Moreover, both models agree on the conclusion that turnover is inefficiently low. In particular, the high mobility of young workers in the US labour market should not be interpreted as an alarming indicator of the inefficiency of this labour market.

Section 2 contains the analysis of the informational model. Having set up the model and constructed the equilibrium, we derive the equilibrium underprovision of training in Theorem 2. Section 3 presents the productivity-enhancement model, and derives the overprovision of training in Theorem 5. Section 4 concludes.

2. An Informational Model

2.1. The Dynamics. There are two firms, indexed by $i \in \{1, 2\}$, and one worker. At each time $t \in [0, +\infty)$, the worker chooses to work for one of the two firms. That firm then either assigns her to one of a range of tasks, indexed by $\lambda \in \Lambda$, or to a training program.

If the worker chooses to work for firm i, and if she is assigned to task λ, then the output of firm i is

$$dy_i = b_i^\lambda(s_i)\,dt + \sigma_i^\lambda\,dW_i;$$

and the output of firm $j \neq i$ is $dy_j = 0$. Here $s_i \in \{0, 1\}$ is the firm-i-specific type of the worker, b_i^λ is a function from $\{0, 1\}$ to $\mathbb{R}$ that determines the worker's productivity as a function of her firm-specific type, σ_i^λ is a strictly positive scalar, and W_i is a firm-i-specific Wiener process. In other words, dy_i is the sum of the fixed effect $b_i^\lambda(s_i)\,dt$ and the firm-i-specific shock $\sigma_i^\lambda\,dW_i$.

The two firms and the worker all observe i, λ, dy_i and dy_j. They can use these observations to update their common beliefs p_i that $s_i = 1$ and p_j that $s_j = 1$. Indeed, it can be shown that dp_i is distributed normally with mean 0 and variance

$$\left(\phi_i^\lambda\,p_i\,(1 - p_i)\right)^2\,dt,$$

where

$$\phi_i^\lambda = \frac{|b_i^\lambda(1) - b_i^\lambda(0)|}{\sigma_i^\lambda}$$

is the signal-to-noise ratio, and that $dp_j = 0$.[9] We interpret $\left(\phi_i^\lambda\,p_i\,(1 - p_i)\right)^2$ as the rate of accumulation of the worker's firm-i-specific human capital at time t.

Now, we want to focus on the difference between the rates of accumulation of firm-specific human capital on the job and on the training program, and not on the differences among the rates of accumulation of firm-specific human capital associated with the various tasks that the worker can undertake when she is on the job. We therefore assume that ϕ_i^λ is independent of λ, and we denote the common value by ϕ_i. This implies that all tasks are equivalent from an informational point of view. Hence, if firm i assigns the worker to the job, then it will assign her to the task with

[9] Cf. Bolton and Harris (1999) and Felli and Harris (1996).

the highest expected output per unit time. Since her expected output per unit time is

$$b_i^\lambda(p_i) = (1 - p_i)\, b_i^\lambda(0) + p_i\, b_i^\lambda(1)$$

in task λ, her expected output per unit time will be

$$M_i(p_i) = \max_{\lambda \in \Lambda} b_i^\lambda(p_i)$$

in the task to which she is actually assigned.

Finally, if the worker chooses to work for firm i, and if she is assigned to the training program, then the output of firm i is

$$dy_i = k_i\, dt,$$

where $k_i < \max\{M_i(0), M_i(1)\}$; and the output of firm $j \neq i$ is $dy_j = 0$. Furthermore, we assume that dp_i is distributed normally with mean 0 and variance

$$\theta_i^{\mathcal{T}}\, (\phi_i\, p_i\, (1 - p_i))^2\, dt,$$

where $\theta_i^{\mathcal{T}} > 1$; and that $dp_j = 0$. In particular, we assume that the worker's output when training is never so high as to imply that training is a dominant strategy, and that the effect of training is to increase the rate of accumulation of firm-i-specific human capital by the factor $\theta_i^{\mathcal{T}} > 1$.

2.2. A Reformulation of the Dynamics. It is helpful to reformulate the dynamics in terms of the expected productivities $m_1 = M_1(p_1)$ and $m_2 = M_2(p_2)$. This change allows us to identify the two key components of training, namely the productivity-enhancement component and the employee-evaluation component. Moreover, it enables us to make direct comparisons between the informational model, analyzed in the present section, and the productivity-enhancement model, analyzed in Section 3 below.

In order to do so, we shall need the following assumptions concerning the productivity schedule M_i:[10]

A1 M_i is symmetric about $p_i = \frac{1}{2}$;

A2 $M_i'' > 0$ on $[0, 1]$.

Assumption A1 is made for expositional convenience. It ensures that we do not need to keep track of whether a given value of m_i arises from a $p_i \in [0, \frac{1}{2})$ or a $p_i \in (\frac{1}{2}, 1]$. Assumption A2 means that information always has a strictly positive value.

[10] We also need a technical assumption, namely that M_i is eight times continuously differentiable. We shall refer to this assumption as Assumption A0. See Appendix A.1.

Now, if the worker chooses to work for firm i, and if she is assigned to the job, then Itô's Lemma tells us that the change in the expected productivity of the worker in firm i is given by the formula

$$dm_i = M_i'(p_i)\,dp_i + \tfrac{1}{2}\,M_i''(p_i)\,dp_i^2.$$

Since dp_i is distributed normally with mean 0 and variance $(\phi_i\,p_i\,(1-p_i))^2\,dt$, this implies that dm_i is distributed normally with mean $\widehat{\mu}_i(p_i)\,dt$ and variance $\widehat{\Psi}_i(p_i)\,dt$, where

$$\widehat{\mu}_i(p_i) = \tfrac{1}{2}M_i''(p_i)\,(\phi_i\,p_i\,(1-p_i))^2$$

and

$$\widehat{\Psi}_i(p_i) = M_i'(p_i)^2\,(\phi_i\,p_i\,(1-p_i))^2.$$

Similarly, the change in the expected productivity of the worker in firm $j \neq i$ is given by the formula

$$dm_j = M_j'(p_j)\,dp_j + \tfrac{1}{2}\,M_j''(p_j)\,dp_j^2.$$

Since $dp_j = 0$, this implies that $dm_j = 0$.

If, on the other hand, the worker chooses to work for firm i, and if she is assigned to the training program, then dp_i is distributed normally with mean 0 and variance $\theta_i^{\mathcal{T}}\,(\phi_j\,p_j\,(1-p_j))^2\,dt$. Hence dm_i is distributed normally with mean $\theta_i^{\mathcal{T}}\,\widehat{\mu}_i(p_i)\,dt$ and variance $\theta_i^{\mathcal{T}}\,\widehat{\Psi}_i(p_i)\,dt$. Similarly, $dp_j = 0$. Hence $dm_j = 0$.

We interpret $\widehat{\mu}_i$ as the productivity-enhancement component of learning by doing, and $\widehat{\Psi}_i$ as the employee-evaluation component of learning by doing. Moreover, we emphasize that the effect of training in our model is to increase both components by the same factor $\theta_i^{\mathcal{T}} > 1$.

Finally, Assumptions A1-A2 imply that M_i is strictly decreasing on $[0, \frac{1}{2}]$, that M_i is strictly increasing on $[\frac{1}{2}, 1]$, and that $M_i(p_i) = M_i(1-p_i)$. Hence m_i lies in the bounded interval $[\underline{m}_i, \overline{m}_i]$, where $\underline{m}_i = M_i(\frac{1}{2})$ and $\overline{m}_i = M_i(1)$. Moreover, there exist functions

$$\mu_i, \Psi_i : [\underline{m}_i, \overline{m}_i] \to [0, +\infty)$$

such that $\widehat{\mu}_i(p_i) = \mu_i(M_i(p_i))$ and $\widehat{\Psi}_i(p_i) = \Psi_i(M_i(p_i))$. Assumptions A1-A2 also imply the following properties of μ_i and Ψ_i:[11]

B1 $\mu_i > 0$ on $[\underline{m}_i, \overline{m}_i)$, and $\mu_i(\overline{m}_i) = 0$;

B2 $\Psi_i > 0$ on $(\underline{m}_i, \overline{m}_i)$, and $\Psi_i(\underline{m}_i) = \Psi_i(\overline{m}_i) = 0$.

Property B1 means that productivity enhancement on the job is strictly positive at all productivity levels except $\overline{m}_i$, where it is 0. Property B2 means that employee evaluation on the job is strictly positive at all productivity levels except $\underline{m}_i$ and $\overline{m}_i$, where it is 0.

[11] More precisely, Assumptions A0-A2 imply Properties B0-B2, where: A0 is the assumption that M_i is eight times continuously differentiable; and B0 is the property that μ_i and Ψ_i are thrice continuously differentiable. See footnote 10 above and Appendix A.1 below.

2.3. The Negotiations between the Firms and the Worker. In order to complete the model, we need to specify the form taken by the negotiations between the two firms and the worker. At each time t, the two firms simultaneously and independently submit an offer to the worker. Firm i's offer consists of a wage $w_i \in \mathbb{R}$ and an assignment $a_i \in \{\mathcal{W}, \mathcal{T}\}$, where $\mathcal{W}$ signifies that the worker will be assigned to the job, and $\mathcal{T}$ signifies that the worker will be assigned to the training program. The worker then chooses one of the two offers, and undertakes the assignment specified in the offer. The employer then pays the worker the wage specified in the offer, and obtains the output associated with the assignment specified in the offer. Finally, the entire cycle recommences.

2.4. Equilibrium. We can now proceed to derive the system of Bellman equations that characterizes equilibrium in our dynamic game.[12] The basic idea is as follows. Suppose that we are given functions

$$U_1, U_2, V : [\underline{m}_1, \overline{m}_1] \times [\underline{m}_2, \overline{m}_2] \to \mathbb{R}.$$

Then U_1, U_2 and V are the value functions of firm 1, firm 2 and the worker in an equilibrium of the dynamic game if and only if the following condition holds: for all states

$$m = (m_1, m_2) \in [\underline{m}_1, \overline{m}_1] \times [\underline{m}_2, \overline{m}_2],$$

if the continuation payoffs of the three players are given by the functions U_1, U_2 and V, then the values

$$U_1(m), U_2(m), V(m)$$

are the payoffs of the three players in an equilibrium of the stage game.

2.5. The Stage Game. Suppose that all three market participants use the same discount function[13]

$$r \exp(-r\,t).$$

Define the expected-output functions $\pi_i^{\mathcal{W}}$ and $\pi_i^{\mathcal{T}}$ by the formulae

$$\begin{aligned} \pi_i^{\mathcal{W}}(m_i) &= m_i, \\ \pi_i^{\mathcal{T}}(m_i) &= k_i. \end{aligned}$$

Suppose further that the current state is m; that the continuation payoffs of the three players are given by the functions U_1, U_2 and V; that the two firms make the

[12] The derivation is the generalization to the case of training of the results of our earlier paper, namely Felli and Harris (1996). The algebra involved in the derivation is somewhat lengthy, but the system of Bellman equations itself is very simple. We therefore divide the derivation into four lemmas and a theorem, collecting as much as possible of the algebra into the proofs. The system of Bellman equations is laid out in the statement of the theorem.

[13] For convenience, we multiply the standard discount function $\exp(-r\,t)$ by the discount rate r. This simplifies the interpretation of many of the formulae below.

offers (w_1, a_1) and (w_2, a_2); and that the worker chooses to work for firm 1. Then the current payoffs of the three players are

$$\begin{aligned}
&r\,(\pi_1^{a_1}(m_1) - w_1)\,dt,\\
&0,\\
&r\,w_1\,dt;
\end{aligned}$$

and their discounted continuation payoffs are

$$\begin{aligned}
&\exp(-r\,dt)\,U_1(m + dm),\\
&\exp(-r\,dt)\,U_2(m + dm),\\
&\exp(-r\,dt)\,V(m + dm).
\end{aligned}$$

Their expected payoffs in the stage game are therefore

$$\begin{aligned}
&\mathrm{E}\big[r\,(\pi_1^{a_1}(m_1) - w_1)\,dt + \exp(-r\,dt)\,U_1(m + dm)\big],\\
&\mathrm{E}\big[\exp(-r\,dt)\,U_2(m + dm)\big],\\
&\mathrm{E}\big[r\,w_1\,dt + \exp(-r\,dt)\,V(m + dm)\big],
\end{aligned}$$

where E denotes expectation conditional on the information available at the start of the current period. In the case where the worker chooses to work for firm 2, the expected payoffs of the three players are analogous.

Now denote partial differentiation with respect to m_i by ∂_i; define the training operator T_i by the formula

$$T_i F(m) = \mu_i(m_i)\,\partial_i F(m) + \tfrac{1}{2}\Psi_i(m_i)\,\partial_i^2 F(m);$$

and put

$$\theta_i^{\mathcal{W}} = 1.$$

Then the expected payoffs in the stage game can be written in a more convenient form.

Lemma 1. *Suppose that the current state is m, and that the continuation values of the three players are given by the functions U_1, U_2 and V. Suppose further that the two firms make the offers (w_1, a_1) and (w_2, a_2), and that the worker chooses to work for firm 1. Then the expected payoffs of firm 1, firm 2 and the worker in the stage game can be written*

$$(1 - r\,dt)\,U_1(m) + r\,dt\,\big(\pi_1^{a_1}(m_1) - w_1 + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_1(m)\big), \tag{1}$$

$$(1 - r\,dt)\,U_2(m) + r\,dt\,\big(\tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_2(m)\big), \tag{2}$$

$$(1 - r\,dt)\,V(m) + r\,dt\,\big(w_1 + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1V(m)\big). \tag{3}$$

The expected payoffs when the worker chooses to work for firm 2 can be written analogously.

Proof. We consider the expected payoff of firm 1. The expected payoffs of firm 2 and the worker can be treated similarly. Itô's Lemma implies that

$$\exp(-r\,dt) = 1 - r\,dt$$

and

$$U_1(m + dm) = U_1(m) + \partial U_1(m) \cdot dm + \tfrac{1}{2}\, dm \cdot \partial^2 U_1(m)\, dm,$$

where

$$\partial U_1(m) \cdot dm = \partial_1 U_1(m)\, dm_1 + \partial_2 U_2(m)\, dm_2$$

and

$$dm \cdot \partial^2 U_1(m)\, dm = \partial_1^2 U_1(m)\, dm_1^2 + 2\,\partial_1 \partial_2 U_1(m)\, dm_1\, dm_2 + \partial_2^2 U_2(m)\, dm_2^2.$$

Hence the discounted continuation payoff of firm 1 is

$$(1 - r\,dt)\, U_1(m) + \partial U_1(m) \cdot dm + \tfrac{1}{2}\, dm \cdot \partial^2 U_1(m)\, dm$$

(dropping terms of order higher than dt, and rearranging)

$$= (1 - r\,dt)\, U_1(m) + \partial_1 U_1(m)\, dm_1 + \tfrac{1}{2}\, \partial_1^2 U_1(m)\, dm_1^2$$

(since $dm_2 = 0$); and the overall payoff of firm 1 is

$$(1 - r\,dt)\, U_1(m) + r\,dt\,(\pi_1^{a_1}(m_1) - w_1) + \partial_1 U_1(m)\, dm_1 + \tfrac{1}{2}\, \partial_1^2 U_1(m)\, dm_1^2.$$

Finally, since dm_1 has mean $\theta_1^{a_1}\, \mu_1(m_1)\, dt$ and variance $\theta_1^{a_1}\, \Psi_1(m_1)\, dt$, the expected payoff of firm 1 is

$$(1 - r\,dt)\, U_1(m) + r\,dt\,(\pi_1^{a_1}(m_1) - w_1 + \tfrac{1}{r}\, \theta_1^{a_1}\, T_1 U_1(m)),$$

as required. ■

2.6. Initial Characterization of Equilibrium. Using Lemma 1, it can be shown that the stage game is strategically equivalent to a new static game with simpler payoffs. Using this equivalence, we obtain the following initial characterization of equilibrium in the dynamic game.

Lemma 2. *U_1, U_2 and V are the value functions of firm 1, firm 2 and the worker in an equilibrium of the dynamic game if and only if, for all states m, the values $U_1(m)$, $U_2(m)$ and $V(m)$ are the payoffs of firm 1, firm 2 and the worker in an equilibrium of the new static game in which: the payoffs when the two firms make the offers (w_1, a_1) and (w_2, a_2), and the worker chooses to work for firm 1, are*

$$\pi_1^{a_1}(m_1) - w_1 + \tfrac{1}{r}\, \theta_1^{a_1}\, T_1 U_1(m), \tag{4}$$

$$\tfrac{1}{r}\, \theta_1^{a_1}\, T_1 U_2(m), \tag{5}$$

$$w_1 + \tfrac{1}{r}\, \theta_1^{a_1}\, T_1 V(m); \tag{6}$$

and the payoffs when the worker chooses to work for firm 2 are analogous.

There are three components to the payoff of firm 1 in the new static game: the expected marginal product $\pi_1^{a_1}(m_1)$ of the worker when she is given the assignment a_1; (minus) the wage w_1; and the discount factor $\frac{1}{r}$ times the acceleration factor $\theta_1^{a_1}$ times the shadow value to firm 1 of human-capital accumulation on the job with firm 1, namely $T_1 U_1(m)$. There is one component to the payoff of firm 2 in the new static game: the discount factor $\frac{1}{r}$ times the acceleration factor $\theta_1^{a_1}$ times the shadow value to firm 2 of human-capital accumulation on the job with firm 1, namely $T_1 U_2(m)$. Finally, there are two components to the payoff of the worker in the new static game: the wage w_1; plus the discount factor $\frac{1}{r}$ times the acceleration factor $\theta_1^{a_1}$ times the shadow value to the worker of human-capital accumulation on the job with firm 1, namely $T_1 V(m)$.

Proof. We have already noted above that U_1, U_2 and V are the value functions of an equilibrium of the dynamic game if and only if, for all states m, if the continuation payoffs of the three players are given by the functions U_1, U_2 and V, then the values $U_1(m)$, $U_2(m)$ and $V(m)$ are the payoffs of an equilibrium of the stage game. Now, Lemma 1 implies that the payoffs (1), (2) and (3) in the stage game are positive affine transformations (with respective intercepts $(1 - r\,dt)\,U_1(m)$, $(1 - r\,dt)\,U_2(m)$ and $(1 - r\,dt)\,V(m)$, and with common slope $r\,dt$) of the payoffs (4), (5) and (6) in the new static game. Moreover the inverses of these transformations map the values $U_1(m)$, $U_2(m)$ and $V(m)$ to themselves. ■

2.7. Analysis of the New Static Game. In the interests of notational simplicity, we suppress the dependence of U_1, U_2 and V on m, the dependence of $\pi_1^{a_1}$ on m_1 and the dependence of $\pi_2^{a_2}$ on m_2. We can then characterize the equilibria of the new static game as follows

Lemma 3. *Suppose that the current state is m. Then the outcome in which the two firms make the offers (w_1, a_1) and (w_2, a_2), and the worker chooses to work for firm 1, is an equilibrium of the new static game iff*

$$w_1 = \pi_2^{a_2} + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2(U_2 + V) - \tfrac{1}{r}\,\theta_1^{a_1}\,T_1(U_2 + V), \tag{7}$$

$$a_1 \in \operatorname*{argmax}_{\tilde{a}_1} \left\{\pi_1^{\tilde{a}_1} + \tfrac{1}{r}\,\theta_1^{\tilde{a}_1}\,T_1(U_1 + V)\right\}, \tag{8}$$

$$w_2 = \pi_2^{a_2} + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2 U_2 - \tfrac{1}{r}\,\theta_1^{a_1}\,T_1 U_2, \tag{9}$$

$$a_2 \in \operatorname*{argmax}_{\tilde{a}_2} \left\{\pi_2^{\tilde{a}_2} + \tfrac{1}{r}\,\theta_2^{\tilde{a}_2}\,T_2(U_2 + V)\right\}, \tag{10}$$

$$1 \in \operatorname*{argmax}_{i} \left\{\pi_i^{a_i} + \tfrac{1}{r}\,\theta_i^{a_i}\,T_i(U_1 + U_2 + V)\right\}. \tag{11}$$

The characterization of the outcome in which the worker works for firm 2 is analogous.

In order to explain these findings, it will be helpful to introduce some terminology. Let us refer to the coalition consisting of firm 1 and the worker as coalition 1, to the coalition consisting of firm 2 and the worker as coalition 2, and to the coalition consisting of all three market participants as the grand coalition.

In this terminology, condition (8) means that a_1 will be chosen to maximize the payoff of coalition 1, and condition (10) means that a_2 will be chosen to maximize the payoff of coalition 2. This is exactly as one would expect: any disagreement between the worker and a firm as to whether the worker should be assigned to the job or to the training program can be resolved by adjusting the wage.

Similarly, condition (11) means that, taking the job-training assignment of the worker within each firm as given, the worker should be allocated between the two firms in such a way as to maximize the payoff of the grand coalition. In other words, the allocation of the worker between the two firms should be constrained efficient. This is again exactly as one would expect. Neither firm can affect the assignment of the worker within the other firm. This assignment therefore acts as a constraint. On the other hand, both firms can affect the worker's choice of employer at the margin, by adjusting the wage. This choice should therefore be efficient.

The wage offer (7) by firm 1 has three components: $\pi_2^{a_2}$ is the output that the worker would have produced if she had worked for firm 2 instead; $\frac{1}{r}\,\theta_2^{a_2}\,T_2(U_2+V)$ is the value to coalition 2 of the human capital that the worker would have acquired if she had worked for firm 2 instead; and $\frac{1}{r}\,\theta_1^{a_1}\,T_1(U_2+V)$ is the value to coalition 2 of the human capital that the worker acquires by doing what she actually does, namely working for firm 1. Overall, then, firm 1 offers to pay the worker the net value to coalition 2 of having the worker switch from working for firm 1 to working for firm 2.

Similarly, the wage offer (9) by firm 2 has three components: $\pi_2^{a_2}$ is the output that the worker would have produced if she had worked for firm 2 instead; $\frac{1}{r}\,\theta_2^{a_2}\,T_2U_2$ is the value to firm 2 of the human capital that the worker would have acquired if she had worked for firm 2 instead; and $\frac{1}{r}\,\theta_1^{a_1}\,T_1U_2$ is the value to firm 2 of the human capital that the worker acquires by doing what she actually does, namely working for firm 1. And, overall, firm 2 offers to pay the worker the net value to itself of having the worker switch from working for firm 1 to working for firm 2.

Proof. Suppose that firm 1's offer (w_1, a_1) is given. Then the value to the worker of firm 2's offer (w_2, a_2) is

$$w_2 + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2V, \tag{12}$$

and firm 2 is willing to make this offer if and only if

$$\pi_2^{a_2} - w_2 + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2U_2 \geq \tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_2, \tag{13}$$

i.e. if and only if the value of having the offer accepted exceeds the value of the fallback option of allowing the worker to work for firm 1. For any given a_2, (12) is maximized given the condition (13) if and only if

$$w_2 = \pi_2^{a_2} + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2U_2 - \tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_2. \tag{14}$$

With this choice of w_2, (12) becomes

$$\pi_2^{a_2} + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2(U_2+V) - \tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_2. \tag{15}$$

This in turn is maximized by choosing

$$a_2 \in \operatorname*{argmax}_{a_2}\left\{\pi_2^{a_2} + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2(U_2+V)\right\}. \tag{16}$$

Suppose now that firm 2's offer (w_2, a_2) is given. Then the value to firm 1 of its own offer (w_1, a_1) is

$$\pi_1^{a_1} - w_1 + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_1, \tag{17}$$

and the worker is willing to accept this offer if and only if

$$w_1 + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1V \geq w_2 + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2V, \tag{18}$$

i.e. if and only if the value of firm 1's offer is at least as high as the value of firm 2's offer. For any given a_1, (17) is maximized given the condition (18) if and only if

$$w_1 = w_2 + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2V - \tfrac{1}{r}\,\theta_1^{a_1}\,T_1V. \tag{19}$$

With this choice of w_1, (17) becomes

$$\pi_1^{a_1} + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1(U_1+V) - w_2 - \tfrac{1}{r}\,\theta_2^{a_2}\,T_2V. \tag{20}$$

This in turn is maximized by choosing

$$a_1 \in \operatorname*{argmax}_{a_1}\left\{\pi_1^{a_1} + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1(U_1+V)\right\}. \tag{21}$$

Finally, firm 1 will be willing to make the offer (w_1, a_1) if and only if

$$\pi_1^{a_1} + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1(U_1+V) - w_2 - \tfrac{1}{r}\,\theta_2^{a_2}\,T_2V \geq \tfrac{1}{r}\,\theta_2^{a_2}\,T_2U_1. \tag{22}$$

Putting these two lines of reasoning together, we see that (w_1, a_1) and (w_2, a_2) form the basis for an equilibrium of the negotiations at time t in which the worker works for firm 1 if and only if (14), (16), (19), (21) and (22) hold. Finally, using (14) to substitute for w_2 in (19) and (22), and rearranging, we see that (w_1, a_1) and (w_2, a_2) form the basis for an equilibrium of the negotiations at time t in which the worker works for firm 1 if and only if (7-11) hold. ■

Lemma 4. *Suppose that the current state is m. Then, in any equilibrium of the new static game in which the workers choose firm 1, the payoffs of the three players are*

$$\pi_1^{a_1} + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1(U_1+U_2+V) - \pi_2^{a_2} - \tfrac{1}{r}\,\theta_2^{a_2}\,T_2(U_2+V), \tag{23}$$

$$\tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_2, \tag{24}$$

$$\pi_2^{a_2} + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2(U_2+V) - \tfrac{1}{r}\,\theta_1^{a_1}\,T_1U_2. \tag{25}$$

The payoffs in any equilibrium in which the worker chooses firm 2 are analogous.

The equilibrium payoff of firm 2 is simply the value to firm 1 of the human capital that the worker obtains by working for firm 1. The equilibrium payoff of the worker is the value to coalition 2 of the human capital that the worker would have acquired if she had worked for firm 2 instead, less the payoff to firm 2. Finally, the equilibrium payoff of firm 1 is the value to the grand coalition of the human capital that the worker obtains by working for firm 1, less the value to coalition 2 of the human capital that the worker would have acquired if she had worked for firm 2 instead.

Proof. Use the formula for the equilibrium wage, namely (7), to substitute for w_1 in the formulae for the payoffs, namely (4), (5) and (6). ■

2.8. Final Characterization of Equilibrium. Let us put $S_1 = U_1 + V$, $S_2 = U_2 + V$ and $S = U_1 + U_2 + V$. Then:

Theorem 1. *S_1, S_2 and S are the value functions of coalition 1, coalition 2 and the grand coalition in an equilibrium of the dynamic game if and only if there exist training policies*

$$A_1, A_2 : [\underline{m}_1, \overline{m}_1] \times [\underline{m}_2, \overline{m}_2] \to \{\mathcal{T}, \mathcal{W}\}$$

and an employment policy

$$I : [\underline{m}_1, \overline{m}_1] \times [\underline{m}_2, \overline{m}_2] \to \{1, 2\}$$

such that, for all states m, we have

$$S_1 = \pi_1^{A_1} + \tfrac{1}{r}\theta_1^{A_1}\, T_1\, S_1, \tag{26}$$

$$S_2 = \pi_2^{A_2} + \tfrac{1}{r}\theta_2^{A_2}\, T_2\, S_2, \tag{27}$$

$$S = \pi_I^{A_I} + \tfrac{1}{r}\theta_I^{A_I}\, T_I\, S \tag{28}$$

and

$$A_1 \in \operatorname*{argmax}_{a_1 \in \{\mathcal{T}, \mathcal{W}\}} \left\{\pi_1^{a_1} + \tfrac{1}{r}\theta_1^{a_1}\, T_1\, S_1\right\}, \tag{29}$$

$$A_2 \in \operatorname*{argmax}_{a_2 \in \{\mathcal{T}, \mathcal{W}\}} \left\{\pi_2^{a_2} + \tfrac{1}{r}\theta_2^{a_2}\, T_2\, S_2\right\}, \tag{30}$$

$$I \in \operatorname*{argmax}_{i \in \{1,2\}} \left\{\pi_i^{A_i} + \tfrac{1}{r}\theta_i^{A_i}\, T_i\, S\right\}. \tag{31}$$

Taken together, (26) and (29) tell us that S_1 is the value function for coalition 1 under autarky, and A_1 is the associated training policy. Similarly, (27) and (30) tell us that S_2 is the value function for coalition 2 under autarky, and A_2 is the associated training policy. Finally, (28) and (31) tell us that, if A_1 and A_2 are given, then S is the value function of the grand coalition, and I is the associated employment policy.

In particular, the model can be solved recursively, and both steps of the recursion involve solving standard optimization problems.

Proof. By definition, S_1, S_2 and S are the value functions of coalition 1, coalition 2 and the grand coalition in an equilibrium of the dynamic game if and only if U_1, U_2 and V are the value functions of firm 1, firm 2 and the worker in an equilibrium of the dynamic game. Moreover Lemma 2 tells us that U_1, U_2 and V are the value functions of an equilibrium of the dynamic game if and only if there exist policies A_1, A_2 and I such that, for all states m, $(A_1(m), A_2(m), I(m))$ is an equilibrium of the new static game in which the players receive the payoffs $U_1(m)$, $U_2(m)$ and $V(m)$.

Now, substituting for U_1, U_2 and V in terms of S_1, S_2 and S in conditions (8), (10) and (11) of Lemma 3, we see that $(A_1(m), A_2(m), I(m))$ is an equilibrium of the new static game if and only if conditions (29), (30) and (31) hold.

Similarly, substituting for U_1, U_2 and V in terms of S_1, S_2 and S in conditions (23), (24) and (25) of Lemma 4, we see that $U_1(m)$, $U_2(m)$ and $V(m)$ are the payoffs in an equilibrium of the new static game if and only if

$$\begin{aligned} U_1(m) &= \pi_1^{a_1}(m) + \tfrac{1}{r}\,\theta_1^{a_1}\,T_1S(m) - \pi_2^{a_2}(m) - \tfrac{1}{r}\,\theta_2^{a_2}\,T_2S_2(m), && (32)\\ U_2(m) &= \tfrac{1}{r}\,\theta_1^{a_1}\,T_1(S-S_1)(m), && (33)\\ V(m) &= \pi_2^{a_2}(m) + \tfrac{1}{r}\,\theta_2^{a_2}\,T_2S_2(m) - \tfrac{1}{r}\,\theta_1^{a_1}\,T_1(S-S_1)(m). && (34) \end{aligned}$$

Moreover, adding equations (32) and (34), adding equations (33) and (34) and adding equations (32), (33) and (34), we see that the system of equations (32), (33) and (34) is equivalent to the system of equations (26), (27) and (28). ■

Finally, we state two immediate consequences of Theorem 1.

Corollary 1. *S_i depends only on m_i.*

In other words, the equilibrium value of coalition 1 does not depend on the expect value of the worker in the alternative match.

Proof. This follows at once from the fact that S_i is the value function for coalition i under autarky. ■

Corollary 2. *The equilibrium wage is S_2 if $I = 1$ and S_1 if $I = 2$.*

In other words, in equilibrium, the worker is paid his outside option, namely the value of the alternative match.

Proof. Suppose for concreteness that $I = 1$. Then the wage offered by firm 1 is

$$\pi_2^{A_2} + \tfrac{1}{r}\,\theta_2^{A_2}\,T_2S_2 - \tfrac{1}{r}\,\theta_1^{A_1}\,T_1S_2$$

(cf. (7)). But

$$\pi_2^{A_2} + \tfrac{1}{r}\,\theta_2^{A_2}\,T_2S_2 = S_2$$

(by (27)), and

$$T_1S_2 = 0$$

(by Corollary 1). ■

2.9. Provision of Training. With the characterization of equilibrium given in Theorem 1 in hand, it is easy to analyze the inefficiency in the market provision of firm-specific training. Indeed, suppose that firm 1 is the employer. Then, as shown in Theorem 1, the bargaining mechanism ensures that the preferences of both firm 1 and the worker as to the work-training choice within firm 1 are fully taken into account. By contrast, the bargaining mechanism does not take into account the preference of firm 2. In order to determine whether training is socially underprovided or socially overprovided, it therefore suffices to determine whether the shadow price to firm 2 of training in firm 1 is positive or negative.

We shall need two lemmas. The first lemma exploits the fact that one possible strategy for the grand coalition is to have the worker work for firm 1 at all times. This strategy yields a payoff of S_1.

Lemma 5. $S \geq S_1$.

Proof. We have

$$S = \max_{i \in \{1,2\}} \left\{ \pi_i^{A_i} + \theta_i^{A_i} T_i S \right\} \geq \pi_1^{A_1} + \theta_1^{A_1} T_1 S.$$

That is, S is a supersolution of the equation

$$S_1 = \pi_1^{A_1} + \theta_1^{A_1} T_1 S_1.$$

It follows that $S \geq S_1$. ■

The second lemma shows that the shadow value to the grand coalition of training by firm 1 exceeds the shadow value to coalition 1 of training by firm 1. This is because the grand coalition uses the information generated by training to improve the allocation of the worker between firms.

Lemma 6. *Suppose that* $I = 1$. *Then* $T_1 S \geq T_1 S_1$.

Proof. We have

$$S_1 = \pi_1^{A_1} + \theta_1^{A_1} T_1 S_1.$$

Also, since $I = 1$, we have

$$S = \pi_1^{A_1} + \theta_1^{A_1} T_1 S.$$

Subtracting the first equation from the second and rearranging, we obtain

$$T_1 S - T_1 S_1 = \frac{S - S_1}{\theta_1^{A_1}}.$$

Finally, $S - S_1 \geq 0$ (by Lemma 5). ■

We can now establish the main result of this section, namely that the shadow value to firm 2 of training in firm 1 is positive.

Theorem 2. *Suppose that $I = 1$. Then $T_1U_2 \geq 0$.*

Indeed, suppose that the worker is currently employed by firm 1. Then there are two possible outcomes. First, the worker may change employers in due course. Second, the worker may remain with firm 1 forever. Switching the worker from the job to the training program does not change the probabilities of these two outcomes, but it does bring forward the time at which the first outcome occurs. Doing so is therefore good for firm 2, which makes profits sooner.

Proof. We have $T_1U_2 = T_1S - T_1S_1$. Moreover Lemma 6 implies that $T_1S - T_1S_1 \geq 0$. ■

We can also show that the shadow value to the worker of training in firm 1 is negative.

Theorem 3. *Suppose that $I = 1$. Then $T_1V \leq 0$.*

Indeed, if the worker is currently employed by firm 1, then switching the worker from the job to the training program brings forward the time at which the worker changes employer. This is bad from the point of view of the worker, since her outside option falls sooner.

Proof. This result is the mirror image of Theorem 2. To see this, consider the shadow value T_1S_2 to coalition 2 of training by firm 1. We have $T_1S_2 = T_1U_2 + T_1V$, by definition of S_2. On the other hand, $T_1S_2 = 0$, since S_2 depends only on m_2. Hence $T_1V + T_1U_2 = 0$ and $T_1V = -T_1U_2$. ■

Finally, it can be shown that $T_1U_1 \geq T_1U_2$.[14] In other words, the shadow value to firm 1 of training by firm 1 exceeds the shadow value to firm 2 of training by firm 1. This result must, however, be interpreted with caution: the preference of firm 1 as to the work-training choice within firm 1 also takes into account the opportunity cost of training, namely the output foregone while the worker is trained.

To summarize, firm 2 likes training, and this implies that training is socially underprovided. By the same token, the worker dislikes training, and must therefore be compensated for undertaking it.

3. A Productivity-Enhancement Model

In the informational model of the previous section, training resulted in a proportionate increase in the productivity-enhancement and employee-evaluation components of learning by doing. In the productivity-enhancement model of this section, training increases the productivity-enhancement component of learning by doing, but leaves the employee-evaluation component unchanged.[15]

[14]Indeed, $T_1U_1 = T_1S - T_1S_2 = T_1S = T_1S_1 + T_1U_2$. Moreover, it can be shown that $T_1S_1 \geq 0$.

[15]Both the informational model and the productivity-enhancement model can be thought of as special cases of a stochastic learning-by-doing model.

3.1. The Model. It is helpful to think of a_i as representing the percentage of the worker's time that she spends on training. We therefore adopt the convention that $\mathcal{W} = 0$ and $\mathcal{T} = 1$. In other words, if $a_i = \mathcal{W}$ then the worker spends 0% of her time training, and if $a_i = \mathcal{T}$ then she spends 100% of her time training. The productivity-enhancement model can then be summarized as follows. At each time t, the two firms simultaneously and independently submit an offer to the worker. Firm i's offer consists of a wage $w_i \in \mathbb{R}$ and an assignment $a_i \in [\mathcal{W}, \mathcal{T}]$. The worker then chooses which offer to accept. If the worker accepts the offer of firm i, then: she undertakes the assignment a_i; she receives the wage w_i; her output is $\pi_i^{a_i}\,dt$; and dm_i is distributed normally with mean $\theta_i(a_i)\,\mu_i\,dt$ and variance $\Psi_i\,dt$. The entire cycle then recommences.

3.2. Assumptions. We assume that Properties B1-B2 hold. In particular, we have $\mu_i \geq 0$. We also assume:

B3 $\mu_i' \leq 0$;

B4 $\mu_i'' \geq 0$.

In other words, the productivity-enhancement component of learning by doing is positive, and decreases with productivity at a decreasing rate. Finally, we assume that:

B5 for all $a_i \in [\mathcal{W}, \mathcal{T}]$, the expected-output function $\pi_i^{a_i} : [\underline{m}_i, \overline{m}_i] \to \mathbb{R}$ is given by the formula

$$\pi_i^{a_i}(m_i) = (1 - a_i)\,m_i + a_i\,k_i;$$

B6 the training function $\theta_i : [\mathcal{W}, \mathcal{T}] \to \mathbb{R}$ is twice continuously differentiable, with $\theta_i(\mathcal{W}) = 1$, $\theta_i' > 0$ and $\theta_i'' < 0$.

In particular, if we think of a_i as representing the fraction of her time that the worker spends on training, then output exhibits constant returns, and productivity enhancement exhibits diminishing returns, with respect to the fraction of time spent training.

3.3. Equilibrium. The system of Bellman equations for the productivity-enhancement model can be derived in exactly the same way as the system of Bellman equations for the informational model, and we have:

Theorem 4. *S_1, S_2 and S are equilibrium value functions for coalition 1, coalition 2 and the grand coalition if and only if there exist training policies*

$$A_1, A_2 : [\underline{m}_1, \overline{m}_1] \times [\underline{m}_2, \overline{m}_2] \to [0, 1]$$

and an employment policy

$$I : [\underline{m}_1, \overline{m}_1] \times [\underline{m}_2, \overline{m}_2] \to \{1, 2\}$$

such that, for all states m,

$$A_1 \in \operatorname*{argmax}_{a_1 \in [0,1]} \left\{ \pi_1^{a_1} + \tfrac{1}{r} \left(\theta_1(a_1)\, \mu_1\, \partial_1 S_1 + \tfrac{1}{2}\, \Psi_1\, \partial_1^2 S_1 \right) \right\}, \tag{35}$$

$$A_2 \in \operatorname*{argmax}_{a_2 \in [0,1]} \left\{ \pi_2^{a_2} + \tfrac{1}{r} \left(\theta_2(a_2)\, \mu_2\, \partial_2 S_2 + \tfrac{1}{2}\, \Psi_2\, \partial_2^2 S_2 \right) \right\}, \tag{36}$$

$$I \in \operatorname*{argmax}_{i \in \{1,2\}} \left\{ \pi_i^{A_i} + \tfrac{1}{r} \left(\theta_i(A_i)\, \mu_i\, \partial_i S + \tfrac{1}{2}\, \Psi_i\, \partial_i^2 S \right) \right\} \tag{37}$$

and

$$S_1 = \pi_1^{A_1} + \tfrac{1}{r} \left(\theta_1(A_1)\, \mu_1\, \partial_1 S_1 + \tfrac{1}{2}\, \Psi_1\, \partial_1^2 S_1 \right), \tag{38}$$

$$S_2 = \pi_2^{A_2} + \tfrac{1}{r} \left(\theta_2(A_2)\, \mu_2\, \partial_2 S_2 + \tfrac{1}{2}\, \Psi_2\, \partial_2^2 S_2 \right), \tag{39}$$

$$S = \pi_I^{A_I} + \tfrac{1}{r} \left(\theta_I(A_I)\, \mu_I\, \partial_I S + \tfrac{1}{2}\, \Psi_I\, \partial_I^2 S \right). \tag{40}$$

In this case, the equilibrium value functions of firm 1, *firm* 2 *and the worker can be recovered from* S_1, S_2 *and* S *using the formulae* $U_1 = S - S_2$, $U_2 = S - S_1$, *and* $V = S_1 + S_2 - S$. ■

The only change as compared with Theorem 1 is that operators of the form

$$\theta_i^{a_i}\, T_i = \theta_i^{a_i} \left(\mu_i\, \partial_i + \tfrac{1}{2}\, \Psi_i\, \partial_i^2 \right)$$

have been replaced with operators of the form

$$\theta_i(a_i)\, \mu_i\, \partial_i + \tfrac{1}{2}\, \Psi_i\, \partial_i^2.$$

In other words: in the informational model, the training factor $\theta_i^{a_i}$ multiplied both the productivity-enhancement operator $\mu_i\, \partial_i$ and the employee-evaluation operator $\frac{1}{2}\, \Psi_i\, \partial_i^2$; but, in the productivity-enhancement model, the training factor $\theta_i(a_i)$ multiplies only the productivity-enhancement operator. This reflects the fact that, in the informational model, training affects both the productivity-enhancement component and the employee-evaluation components of learning by doing, whereas, in the productivity-enhancement model, training affects only the productivity-enhancement component of learning by doing.

3.4. Attitudes to Productivity Enhancement. The bargaining mechanism in the productivity-enhancement model is the same as the bargaining mechanism in the informational model. In order to determine whether training is socially underprovided or socially overprovided, it therefore again suffices to determine whether the preference of firm 2 as to the work-training choice within firm 1 is positive or negative.

We shall need two lemmas. The first lemma shows that the shadow value to coalition 1 of productivity enhancement in firm 1 is positive.

Lemma 7. $\partial_1 S_1 \geq 0$.

Although the proof of this lemma is rather long, the underlying rationale for the result is simple: the output $\pi_i^{\mathcal{W}} = m_1$ when the worker works and the output $\pi_i^{\mathcal{T}} = k_1$ when the worker trains are both non-decreasing in m_1. The complications in the proof derive from the fact that we have to take into account the dependence of μ_1 and Ψ_1 on m_1 as well.

Proof. The Bellman equation for S_1 is:

$$S_1 = \max_{a_1 \in [0,1]} \left\{ \pi_1^{a_1} + \tfrac{1}{r} \left(\theta_1(a_1)\, \mu_1\, \partial_1 S_1 + \tfrac{1}{2}\, \Psi_1\, \partial_1^2 S_1 \right) \right\}.$$

Differentiating with respect to m_1 and using the envelope principle, we obtain the equation of variations for $\partial_1 S_1$:

$$\partial_1 S_1 = \partial_1 \pi_1^{A_1} + \tfrac{1}{r}\, \theta_1(A_1) \left(\mu_1'\, \partial_1 S_1 + \mu_1\, \partial_1^2 S_1 \right) + \tfrac{1}{2}\, \tfrac{1}{r} \left(\Psi_1'\, \partial_1^2 S_1 + \Psi_1\, \partial_1^3 S_1 \right).$$

Rearranging, we obtain:

$$\left(1 - \tfrac{1}{r}\, \theta_1(A_1)\, \mu_1' \right) \partial_1 S_1 = \partial_1 \pi_1^{A_1} + \tfrac{1}{r} \left(\theta_1(A_1)\, \mu_1 + \tfrac{1}{2}\, \Psi_1' \right) \partial_1^2 S_1 + \tfrac{1}{2}\, \tfrac{1}{r}\, \Psi_1\, \partial_1^3 S_1$$

In other words, for all $\widehat{m}_1 \in [\underline{m}_1, \overline{m}_1]$, we have the representation

$$\partial_1 S_1(\widehat{m}_1) = \mathrm{E} \left[\int_0^{+\infty} \exp \left(- \int_0^t \beta(m_1(s))\, ds \right) \gamma(m_1(t))\, dt \right],$$

where $m_1(0) = \widehat{m}_1$, $dm_1(t)$ is distributed normally with mean $\zeta(m_1(t))\, dt$ and variance $\eta(m_1(t))\, dt$, and

$$\begin{aligned} \beta &= 1 - \tfrac{1}{r}\, \theta_1(A_1)\, \mu_1', \\ \gamma &= \partial_1 \pi_1^{A_1}, \\ \zeta &= \tfrac{1}{r} \left(\theta_1(A_1)\, \mu_1 + \tfrac{1}{2}\, \Psi_1' \right), \\ \eta &= \tfrac{1}{r}\, \Psi_1. \end{aligned}$$

Now, the discount rate β is bounded. Hence the cumulative discount rate

$$\int_0^t \beta(m_1(s))\, ds$$

is well defined and finite. Hence the discount factor

$$\exp \left(- \int_0^t \beta(m_1(s))\, ds \right)$$

is well defined, finite and strictly positive. Moreover the flow payoff

$$\gamma = \partial_1 \pi_1^{A_1} = (1 - A_1)\, \partial_1 \pi_i^{\mathcal{W}} + A_1\, \partial_1 \pi_i^{\mathcal{T}} = 1 - A_1 \geq 0.$$

Hence the integrand

$$\exp\left(-\int_0^t \beta(m_1(s))\,ds\right)\gamma(m_1(t))$$

is well defined, finite and positive. Hence the representation is well defined in the extended sense and positive. (It could in principle be $+\infty$.) Hence $\partial_1 S_1(\widehat{m}_1) \geq 0$. ■

The second lemma is technical. Put $\widetilde{\theta}_i = \theta_i(A_i)$. Then:

Lemma 8. $1 - \frac{1}{r}\big(\widetilde{\theta}_1\,\mu_1\big)'$ *is bounded above.*

Proof. See Appendix A.2. ■

Remark 1. *The proof of Lemma 8 exploits Properties B3 and B4. It also contains material of independent interest. For example, it shows that* $0 < \partial_1 S_1 \leq 1$ *and* $0 \leq \partial_1^2 S_1 < +\infty$.

We can now prove the main result of this section, namely that the shadow value to firm 2 of productivity enhancement in firm 1 is negative.

Theorem 5. $\partial_1 U_2 \leq 0$.

Indeed, if the worker is currently employed by firm 1, then raising the productivity of the worker in firm 1 both reduces the probability with which the worker changes employer and delays the time at which the change occurs. This is bad from the point of view of firm 2, which is now less likely to make any profit and, if it does make a profit, does so later.

However, the proof suggests an alternative intuition: an increase in the productivity of the worker in firm 1 persists forever. Hence, following such an increase, firm 2 will always have to bid higher to secure the services of the worker. In some cases, this means that firm 2 will no longer be able to afford to secure the services of the worker. In other cases, it means that firm 2 can still secure them, but only at a higher price. Either way, firm 2 is worse off.

Proof. Put $\widetilde{\pi}_i = \pi_i^{A_i}$. Then the Bellman equation for S takes the form

$$S = \max_{i\in\{1,2\}}\left\{\widetilde{\pi}_i + \tfrac{1}{r}\left(\widetilde{\theta}_i\,\mu_i\,\partial_i S + \tfrac{1}{2}\,\Psi_i\,\partial_i^2 S\right)\right\}.$$

Differentiating with respect to m_1 and using the envelope principle, we obtain

$$\partial_1 S = \widetilde{\pi}_1' + \tfrac{1}{r}\left(\tfrac{1}{r}\big(\widetilde{\theta}_1\,\mu_1\big)'\partial_1 S + \widetilde{\theta}_1\,\mu_1\,\partial_1^2 S\right) + \tfrac{1}{2}\,\tfrac{1}{r}\left(\Psi_1'\,\partial_1^2 S + \Psi_1\,\partial_1^3 S\right) \tag{41}$$

if $I = 1$ and

$$\partial_1 S = \tfrac{1}{r}\left(\widetilde{\theta}_2\,\mu_2\,\partial_1\partial_2 S + \tfrac{1}{2}\,\Psi_2\,\partial_1\partial_2^2 S\right) \tag{42}$$

if $I = 2$. Similarly, the Bellman equation for S_1 takes the form

$$S_1 = \widetilde{\pi}_1 + \tfrac{1}{r}\left(\widetilde{\theta}_1\,\mu_1\,\partial_1 S_1 + \tfrac{1}{2}\,\Psi_1\,\partial_1^2 S_1\right).$$

Differentiating with respect to m_1 and using the envelope principle, we obtain

$$\partial_1 S_1 = \widetilde{\pi}_1' + \tfrac{1}{r}\Big(\big(\widetilde{\theta}_1\,\mu_1\big)'\partial_1 S_1 + \widetilde{\theta}_1\,\mu_1\,\partial_1^2 S_1\Big) + \tfrac{1}{2}\,\tfrac{1}{r}\Big(\Psi_1'\,\partial_1^2 S_1 + \Psi_1\,\partial_1^3 S_1\Big). \tag{43}$$

On the other hand, since S_1 is independent of m_2, we have $\partial_2^2 S_1 = \partial_2 S_1 = 0$. Hence

$$\partial_1 S_1 = \partial_1 S_1 + \tfrac{1}{r}\Big(\widetilde{\theta}_2\,\mu_2\,\partial_1\partial_2 S_1 + \tfrac{1}{2}\,\Psi_2\,\partial_1\partial_2^2 S_1\Big). \tag{44}$$

Subtracting (43) from (41), noting that $S - S_1 = U_2$ and rearranging, we obtain

$$\Big(1 - \tfrac{1}{r}\big(\widetilde{\theta}_1\,\mu_1\big)'\Big)\,\partial_1 U_2 = \tfrac{1}{r}\Big(\widetilde{\theta}_1\,\mu_1 + \tfrac{1}{2}\,\Psi_1'\Big)\,\partial_1^2 U_2 + \tfrac{1}{2}\,\tfrac{1}{r}\,\Psi_1\,\partial_1^3 U_2 \tag{45}$$

if $I = 1$. Similarly, subtracting (44) from (42), noting that $S - S_1 = U_2$ and rearranging, we obtain

$$\partial_1 U_2 = -\partial_1 S_1 + \tfrac{1}{r}\Big(\widetilde{\theta}_2\,\mu_2\,\partial_2\partial_1 U_2 + \tfrac{1}{2}\,\Psi_2\,\partial_2^2\partial_1 U_2\Big) \tag{46}$$

if $I = 2$. In other words, for all $\widehat{m} \in [\underline{m}_1, \overline{m}_1] \times [\underline{m}_2, \overline{m}_2]$, we have the representation

$$\partial_1 U_2(\widehat{m}) = \mathrm{E}\left[\int_0^{+\infty} \exp\left(-\int_0^t \beta(m(s))\,ds\right)\gamma(m(t))\,dt\right],$$

where $m(0) = \widehat{m}$, $dm(t)$ is distributed normally with mean $\zeta(m(t))\,dt$ and variance $\eta(m(t))\,dt$, and

$$\begin{aligned}
\beta &= 1 - \tfrac{1}{r}\big(\widetilde{\theta}_1\,\mu_1\big)', \\
\gamma &= 0, \\
\zeta &= \tfrac{1}{r}\Big(\widetilde{\theta}_1\,\mu_1 + \tfrac{1}{2}\,\Psi_1'\Big), \\
\eta &= \tfrac{1}{r}\,\Psi_1
\end{aligned}$$

if $I = 1$ and

$$\begin{aligned}
\beta &= 1, \\
\gamma &= -\partial_1 S_1, \\
\zeta &= \tfrac{1}{r}\,\widetilde{\theta}_2\,\mu_2, \\
\eta &= \tfrac{1}{r}\,\Psi_2
\end{aligned}$$

if $I = 2$.

Now, Lemma 8 implies that β is bounded above. Hence the cumulative discount rate

$$\int_0^t \beta(m(s))\,ds$$

is well defined in the extended sense. (It could in principle be $-\infty$; but if it reaches $-\infty$, it remains there.) Hence the discount factor

$$\exp\left(-\int_0^t \beta(m(s))\,ds\right)$$

is well defined in the extended sense. (It could in principle be $+\infty$; but if it reaches $+\infty$, it remains there.) Moreover $\gamma \le 0$. Hence the integrand

$$\exp\left(-\int_0^t \beta(m(s))\,ds\right)\gamma(m(t))$$

is negative. (It could in principle be $-\infty$.) Hence the representation is well defined in the extended sense. (It could in principle be $-\infty$.) Hence $\partial_1 U_2(\widehat{m}) \le 0$. ■

We can also show that the shadow value to the worker of productivity enhancement in firm 1 is positive.

Theorem 6. $\partial_1 V \ge 0$.

Indeed, if the worker is currently employed by firm 1, then raising the productivity of the worker in firm 1 both reduces the probability with which the worker changes employer and delays the time at which the change occurs. This is good from the point of view of the worker, who is now less likely to experience any fall in her outside option and, if she does experience a fall, does so later.

However, the proof suggests an alternative intuition: an increase in the productivity of the worker in firm 1 persists forever. Hence, following such an increase, firm 2 will always have to bid higher to secure the services of the worker. In some cases, this means that firm 2 will no longer be able to afford to secure the services of the worker. In other cases, it means that firm 2 can still secure them, but only at a higher price. Either way, the worker is better off.[16]

Proof. This result is the mirror image of Theorem 5. Consider the shadow value $\partial_1 S_2$ to coalition 2 of productivity enhancement in firm 1. We have $\partial_1 S_2 = \partial_1 U_2 + \partial_1 V$, by definition of S_2. Moreover $\partial_1 S_2 = 0$, since S_2 depends only on m_2. Hence $\partial_1 V + \partial_1 U_2 = 0$ and $\partial_1 V = -\partial_1 U_2$. ■

Finally, it can be shown that $\partial_1 U_1 \ge \partial_1 U_2$.[17] In other words, the shadow value to firm 1 of productivity enhancement in firm 1 exceeds the shadow value to firm 2 of productivity enhancement in firm 1.

[16]It is possible to make a parallel between the result derived here and the analysis of general training in the presence of market power. In the productivity-enhancement model, where training is firm specific, a worker is able to capture part for the returns from training if the worker is the scarce factor of production. This implies that the worker is willing to pay for firm-specific training. In a world in which training is general, the employer is able to capture part of the returns from training if the employer has labour-market power. The employer is then willing to pay for general training (Stevens 1994a, Stevens 1994b, Acemoglu and Pischke 1998, Acemoglu and Pischke 1999).

[17]Indeed, $\partial_1 U_1 = \partial_1 S - \partial_1 S_2 = \partial_1 S = \partial_1 S_1 + \partial_1 U_2$. Moreover $\partial_1 S_1 \ge 0$, by Lemma 7.

To summarize, firm 2 dislikes productivity enhancement, and this implies that training is socially overprovided. By the same token, the worker likes productivity enhancement, and is therefore willing to take a wage cut in order to obtain training.

3.5. Attitudes to Employee Evaluation. These results on training obtained in the preceding subsection are the opposite of those obtained for the informational model. In order to understand the sharp contrast between the two sets of results, it is helpful to investigate the attitudes of the three market participants to employee evaluation in firm 1. We begin by showing that the shadow price to firm 2 of employee evaluation in firm 1 is positive.

Theorem 7. *Suppose that $I = 1$. Then $\partial_1^2 U_2 \geq 0$.*

Indeed, if the worker is currently employed by firm 1, then increasing employee evaluation in firm 1 both increases the probability with which the worker changes employer and brings forward the time at which the change occurs. This is good from the point of view of firm 2, which is now more likely to make a profit and, when it does make a profit, does so sooner.

Proof. The Bellman equation for S_1 yields

$$S_1 = \pi_1^{A_1} + \tfrac{1}{r}\left(\theta_1(A_1)\,\mu_1\,\partial_1 S_1 + \tfrac{1}{2}\,\Psi_1\,\partial_1^2 S_1\right);$$

and, if $I = 1$, the Bellman equation for S yields

$$S = \pi_1^{A_1} + \tfrac{1}{r}\left(\theta_1(A_1)\,\mu_1\,\partial_1 S + \tfrac{1}{2}\,\Psi_1\,\partial_1^2 S\right).$$

Subtracting the first equation from the second and rearranging, we obtain

$$\partial_1^2 U_2 = \partial_1^2 (S - S_1) = \frac{2}{\Psi_1}\left(r\,(S - S_1) - \theta_1(A_1)\,\mu_1\,\partial_1 (S - S_1)\right).$$

Finally, the same argument that we used in the proof of Lemma 5 shows that $S - S_1 \geq 0$, and Theorem 5 shows that $\partial_1 (S - S_1) = \partial_1 U_2 \leq 0$. Hence $\partial_1^2 U_2 \geq 0$. ■

Next, we show that the shadow value to the worker of employee evaluation in firm 1 is negative.

Theorem 8. $\partial_1^2 V \leq 0$.

Indeed, if the worker is currently employed by firm 1, then increasing employee evaluation in firm 1 both increases the probability with which the worker changes employer and brings forward the time at which the change occurs. This is bad from the point of view of the worker, who is now more likely to experience a fall in her outside option and, when she does experience a fall, does so sooner.

Proof. This result is the mirror image of Theorem 7. Consider the shadow value $\partial_1^2 S_2$ to coalition 2 of employee evaluation in firm 1. We have $\partial_1^2 S_2 = \partial_1^2 U_2 +$

$\partial_1^2 V$, by definition of S_2. Moreover $\partial_1^2 S_2 = 0$, since S_2 depends only on m_2. Hence $\partial_1^2 V + \partial_1^2 U_2 = 0$ and $\partial_1^2 V = -\partial_1^2 U_2$. ■

Finally, it can be shown that $\partial_1^2 U_1 \geq \partial_1^2 U_2$.[18] In other words, the shadow value to firm 1 of employee evaluation in firm 1 exceeds the shadow value to firm 2 of employee evaluation in firm 1.

To summarize, firm 2 likes employee evaluation. By the same token, the worker dislikes employee evaluation. This leads to the following qualitative suggestion: in a model of training that involves a substantial element of employee evaluation, the attitudes of firm 2 and the worker to training will be the reverse of what they are in the productivity-enhancement model. The informational model of Section 2 supports this suggestion. Indeed, it leads to the following sharper suggestion: in a model of training in which the effect of training is to increase employee evaluation by a factor at least as great as the factor by which it increases productivity enhancement, the attitudes of firm 2 and the worker to training will be the reverse of what they are in the productivity-enhancement model.

4. Conclusion

There are good reasons to expect the market to provide workers with the correct incentives to invest in general training.[19] The main question in the context of general training is therefore whether workers are in a position to respond to those incentives. They may not be. For example, the most efficient time for a worker to make an investment in general human capital may be at the outset of her career, when she may not have any financial resources of her own. If so, then she will need to borrow the resources required to make the investment. Unfortunately, financial lenders may not be willing to provide these resources, if the only collateral the worker can offer is her own future labour. Moreover employers may not be willing to underwrite the investment either, because employment law may prevent them from writing the long-term contract necessary to recoup their investment. There may therefore be a case for government intervention, if only partially to reinstate the market that government itself has eliminated. For example, the government could offer a training loan to the worker, and use its powers of taxation to recoup the loan once the worker returns to productive employment. There does not, however, appear to be any case for the direct regulation of employers.

On the other hand, as the analysis of the current paper has shown, there is no reason to expect the market to provide firms and workers with the correct incentives to invest in firm-specific training. The main question in this context is therefore whether there is a case for government intervention to rectify those incentives.

There are at least two obstacles to such intervention. First, in order to determine whether firm-specific training is underprovided or overprovided, it is necessary to

[18] Indeed, $\partial_1^2 U_1 = \partial_1^2 S - \partial_1^2 S_2 = \partial_1^2 S = \partial_1^2 S_1 + \partial_1^2 U_2$. Moreover $\partial_1^2 S_1 \geq 0$ by Lemma 13 in the Appendix.

[19] See Becker (1993).

identify the productivity-enhancement and employee-evaluation components of training. This is a difficult problem. For one thing, the standard assumption – namely that a worker's productivity can be identified with her wage – is not correct in a context in which human capital is firm specific.[20] Moreover, identifying the mean and variance of changes in a worker's productivity is even more challenging than identifying her productivity as such.

Secondly, even if it is possible to identify the two components of specific training, any policy intervention designed to rectify the inefficiency in the provision of such training is likely to encounter opposition from workers. Indeed, as we show in Theorem 3 above, if the employee-evaluation component of training predominates, then: although training is *already underprovided*, workers would prefer to see *less* training. They will therefore resist a policy designed to increase training. Similarly, as we show in Theorem 6 above, if the productivity-enhancement component of training predominates, then: although training is *already overprovided*, workers would prefer to see *more* training. They will therefore resist a policy designed to decrease training.

[20]See Becker (1993), Felli and Harris (1996) and Postel-Vinay and Robin (2002). Becker (1993) points out that the wage of a worker with firm-specific human capital will typically not be equal to her productivity. It might be less than her productivity, because she might be unable to capture the full return to her human capital; and it might be more than her productivity, because her employer might attach option value to retaining her. (Cf. footnote 8 above.) Felli and Harris (1996) present an equilibrium model in which the worker is paid her dynamic outside option. Other things being equal, this is less than her productivity with her current employer when her productivity with her current employer is significantly higher than her productivity with the other firm; and it is greater than her productivity with her current employer when her productivity with her current employer is comparable to her productivity with the other firm. Corollaries 1 and 2 above show that the same is true in the model of the current paper. Postel-Vinay and Robin (2002) explicitly address the problem of identifying the worker's productivity in a world where the worker's wage differs from her marginal productivity. In particular, they analyze and estimate a model with heterogeneous workers and firms. Workers do not accumulate any human capital during their life cycle but progressively learn about their alternative job opportunities while on the job. Whenever a worker receives a new offer she might either accept it or renegotiate her wage with her current employer. In equilibrium, the worker's wage is then her outside option given the sequence of offers she has received up to that point in time. The wage is then a lower bound to the worker's productivity in his current employment. Using the identifying restrictions imposed by their theoretical model, Postel-Vinay and Robin estimate a structural model and explicitly identify the worker's productivity within a match.

A. Appendix

A.1. Assumptions A imply Properties B. In this appendix, we show that Assumptions A0-A2 imply Properties B0-B2, where – as mentioned in footnote 10 – Assumption A0 takes the form:

A0 M_i is eight times continuously differentiable;

and – as mentioned in footnote 11 – Property B0 takes the form:

B0 μ_i and Ψ_i are thrice continuously differentiable.

To this end, let $\Phi_i : [\frac{1}{2}, 1] \to \mathbb{R}$ be given by the formula

$$\Phi_i(p_i) = (\phi_i\, p_i\, (1 - p_i))^2;$$

and let $P_i : [\underline{m}_i, \overline{m}_i] \to [\frac{1}{2}, 1]$ be given implicitly by the formula

$$M_i(P_i(m_i)) = m_i.$$

Then

$$\begin{aligned} \mu_i(m_i) &= \tfrac{1}{2} M_i''(P_i(m_i))\, \Phi_i(P_i(m_i)), \\ \Psi_i(m_i) &= \left(M_i'(P_i(m_i))\right)^2\, \Phi_i(P_i(m_i)). \end{aligned}$$

Hence: $\mu_i > 0$ on $[\underline{m}_i, \overline{m}_i)$ (since $M_i'', \Phi_i > 0$ on $[\frac{1}{2}, 1)$); $\mu_i(\overline{m}_i) = 0$ (since $\Phi_i(1) = 0$); $\Psi_i > 0$ on $(\underline{m}_i, \overline{m}_i)$ (since $M_i', \Phi_i > 0$ on $(\frac{1}{2}, 1)$); $\Psi_i(\underline{m}_i) = 0$ (since $M_i'(\frac{1}{2}) = 0$); and $\Psi_i(\overline{m}_i) = 0$ (since $\Phi_i(1) = 0$).

A full proof that μ_i and Ψ_i are three times continuously differentiable on $[\underline{m}_i, \overline{m}_i]$ would take up too much space, so we confine ourselves to proving the result in the case of the first derivatives. Note first that P_i satisfies

$$M_i(P_i(m_i)) = m_i.$$

Hence

$$M_i'\, P_i' = 1.$$

Secondly,

$$\mu_i(m_i) = \tfrac{1}{2} M_i''(P_i(m_i)) \Phi_i(P_i(m_i)).$$

Hence

$$\mu_i' = \tfrac{1}{2}\,(M_i'''\,\Phi_i + M_i''\,\Phi_i')\, P_i' = \frac{M_i'''\,\Phi_i + M_i''\,\Phi_i'}{2\, M_i'}.$$

But M_i and Φ_i are both symmetric about $p = \frac{1}{2}$. Hence $M_i'(\frac{1}{2}) = M_i'''(\frac{1}{2}) = \Phi_i'(\frac{1}{2}) = 0$, and

$$\begin{aligned} M_i' &= \left(p - \tfrac{1}{2}\right) M_i''(\tfrac{1}{2}) + \mathrm{o}\left(p - \tfrac{1}{2}\right), \\ M_i'' &= M_i''(\tfrac{1}{2}) + \mathrm{o}\left(p - \tfrac{1}{2}\right), \\ M_i''' &= \left(p - \tfrac{1}{2}\right) M_i''''(\tfrac{1}{2}) + \mathrm{o}\left(p - \tfrac{1}{2}\right), \\ \Phi_i &= \Phi_i(\tfrac{1}{2}) + \mathrm{o}\left(p - \tfrac{1}{2}\right), \\ \Phi_i' &= \left(p - \tfrac{1}{2}\right) \Phi_i''(\tfrac{1}{2}) + \mathrm{o}\left(p - \tfrac{1}{2}\right). \end{aligned}$$

Hence

$$\mu_i' = \frac{M_i''''(\frac{1}{2})\,\Phi_i(\frac{1}{2}) + M_i''(\frac{1}{2})\,\Phi_i''(\frac{1}{2})}{2\,M_i''(\frac{1}{2})} + \mathrm{o}\big(p - \tfrac{1}{2}\big)\,.$$

That is, μ_i' is continuous at $p = \frac{1}{2}$. Thirdly,

$$\Psi_i(m_i) = (M_i'(P_i(m_i)))^2\,\Phi_i(P_i(m_i)).$$

Hence

$$\Psi_i' = \Big(2\,M_i'\,M_i''\,\Phi_i + (M_i')^2\,\Phi_i'\Big)\,P_i' = \frac{2\,M_i'\,M_i''\,\Phi_i + (M_i')^2\,\Phi_i'}{M_i'} = 2\,M_i''\,\Phi_i + M_i'\,\Phi_i'.$$

Hence, in the light of the expansions for M_i', M_i'', Φ_i and Φ_i' given above,

$$\Psi_i' = 2\,M_i''(\tfrac{1}{2})\Phi_i(\tfrac{1}{2}) + \mathrm{o}\big(p - \tfrac{1}{2}\big)\,.$$

Hence Ψ_i' is continuous at $p = \frac{1}{2}$.

The second and third derivatives can be handled in the same way. Just as the proof that the first derivative of μ_i is continuous at $p = \frac{1}{2}$ depends on the existence of four continuous derivatives for M_i, so the proof that the second (third) derivative of μ_i is continuous at $p = \frac{1}{2}$ depends on the existence of six (eight) continuous derivatives for M_i. Similarly, just as the proof that the first derivative of Ψ_i is continuous at $p = \frac{1}{2}$ depends on the existence of two continuous derivatives for M_i, so the proof that the second (third) derivative of Ψ_i is continuous at $p = \frac{1}{2}$ depends on the existence of four (six) continuous derivatives for M_i.

A.2. Proof of Lemma 8. Lemma 8 actually subsumes a substantial amount of material, some of which is of interest in its own right. We therefore divide the proof into lemmas. We begin by showing that S_1 is bounded.

Lemma 9. $\underline{m}_1 \le S_1 \le \overline{m}_1$.

Indeed: coalition 1 can ensure that its flow payoff is at least $\underline{m}_1$ at all times, by having the worker work at all times; and it can never achieve a flow payoff greater than $\overline{m}_1$, since $k_1 < \overline{m}_1$.

Proof. We have

$$\max_{a_1\in[0,1]} \big\{\pi_1^{a_1} + \tfrac{1}{r}\left(\theta_1(a_1)\,\mu_1\,\partial_1\overline{m}_1 + \tfrac{1}{2}\,\Psi_1\,\partial_1^2\overline{m}_1\right)\big\} = \max_{a_1\in[0,1]} \{\pi_1^{a_1}\}$$

(since $\partial_1^2\overline{m}_1 = \partial_1\overline{m}_1 = 0$)

$$= \max\{k_1, m_1\} \le \max\{k_1, \overline{m}_1\} \le \overline{m}_1$$

(since $k_1 < \overline{m}_1$). That is, $\overline{m}_1$ is a supersolution of the Bellman equation for S_1. It follows that $S_1 \le \overline{m}_1$. Similarly,

$$\max_{a_1\in[0,1]} \big\{\pi_1^{a_1} + \tfrac{1}{r}\left(\theta_1(a_1)\,\mu_1\,\partial_1\underline{m}_1 + \tfrac{1}{2}\,\Psi_1\,\partial_1^2\underline{m}_1\right)\big\} = \max_{a_1\in[0,1]} \{\pi_1^{a_1}\}$$

(since $\partial_1^2 \underline{m}_1 = \partial_1 \underline{m}_1 = 0$)

$$= \max\{k_1, m_1\} \geq \max\{k_1, \underline{m}_1\} \geq \underline{m}_1.$$

That is, $\underline{m}_1$ is a subsolution of the Bellman equation for S_1. It follows that $S_1 \geq \underline{m}_1$. ■

Next, we exploit Property B3 to show that $\partial_1 S_1 \leq 1$.

Lemma 10. $\partial_1 S_1 \leq 1$.

Indeed, the output m_1 when the worker works and the output k_1 when the worker trains both increase with m_1 at rate at most 1. Moreover μ_1 is decreasing in m_1.

Proof. In the proof of Lemma 7, we showed that $\partial_1 S_1(\widehat{m}_1)$ could be represented as an expected discounted present value with discount rate β, flow payoff γ, mean per unit time ζ and variance per unit time η. Now

$$\gamma = \partial_1 \pi_1^{A_1} = (1 - A_1)\, \partial_1 \pi_i^{\mathcal{W}} + A_1\, \partial_1 \pi_i^{\mathcal{T}} = 1 - A_1 \leq 1$$

and

$$\beta = 1 - \tfrac{1}{r}\, \theta_1(A_1)\, \mu_1' \geq 1.$$

Hence

$$\int_0^{+\infty} \exp\left(-\int_0^t \beta(m_1(s))\, ds\right) \gamma(m_1(t))\, dt \leq 1.$$

Hence $\partial_1 S_1(\widehat{m}_1) \leq 1$. ■

Next, we show that the worker works full time when m_1 is close to its upper bound $\overline{m}_1$.

Lemma 11. *There exists $l_1 \in (k_1, \overline{m}_1)$ such that $A_1 = 0$ on $[l_1, \overline{m}_1]$.*

Indeed, when m_1 is close to $\overline{m}_1$: the benefits of training are small, because there is little scope for increasing productivity; and the opportunity cost of training, namely the output $m_1 - k_1$ foregone, is positive. Hence the worker works.

Proof. The optimality condition for A_1 takes the form

$$A_1 \in \operatorname*{argmax}_{a_1 \in [0,1]} \left\{-(m_1 - k_1)\, a_1 + \tfrac{1}{r}\, \theta_1(a_1)\, \mu_1\, \partial_1 S_1\right\}.$$

Now, for $m_1 \in \left[\frac{1}{2}(k_1 + \overline{m}_1), \overline{m}_1\right]$, we have $m_1 - k_1 \geq \frac{1}{2}(\overline{m}_1 - k_1) > 0$. Moreover $\partial_1 S_1$ bounded (by Lemmas 7 and 10), $\mu_1(\overline{m}_1) = 0$ and μ_1 is continuous. Hence there exists $l_1 \in (k_1, \overline{m}_1)$ such that $\frac{1}{r}\, \theta_1'(\mathcal{W})\, \mu_1\, \partial_1 S_1 < m_1 - k_1$ for $m_1 \in [l_1, \overline{m}_1]$. For such m_1, $A_1(m_1) = 0$. ■

Next, we show that the value function of coalition 1 is strictly increasing in m_1.

Lemma 12. $\partial_1 S_1 > 0$.

Indeed, increasing m_1 increases the flow payoff of coalition 1 in any state in which the worker works. Moreover, we have just seen that the worker works whenever m_1 is close to $\overline{m}_1$.

Proof. In the proof of Lemma 7, we showed that $\partial_1 S_1(\widehat{m}_1)$ could be represented as an expected discounted present value with discount rate β, flow payoff γ, mean per unit time ζ and variance per unit time η. Now:

1. We have $\mu_1(\underline{m}_1) > 0$ and $\Psi_1(\underline{m}_1) = 0$. Moreover, as shown in Appendix A.1, $\Psi_1'(\underline{m}_1) = 2\, M_i''(\frac{1}{2})\, \Phi_i(\frac{1}{2}) > 0$. Hence

$$\begin{aligned} \zeta(\underline{m}_1) &= \tfrac{1}{r}\left(\theta_1(A_1(\underline{m}_1))\, \mu_1(\underline{m}_1) + \tfrac{1}{2}\, \Psi_1'(\underline{m}_1)\right) > 0, \\ \eta(\underline{m}_1) &= \tfrac{1}{r}\, \Psi_1(\underline{m}_1) = 0. \end{aligned}$$

   In other words, m_1 will always move from $\underline{m}_1$ into the interior of $[\underline{m}_1, \overline{m}_1]$.

2. For all $m_1 \in (\underline{m}_1, \overline{m}_1)$, we have

$$\eta(m_1) = \tfrac{1}{r}\, \Psi_1(m_1) > 0.$$

   In other words, m_1 will never come to rest in the interior of $[\underline{m}_1, \overline{m}_1]$.

3. We have $\beta = 1 - \frac{1}{r}\, \theta_1(A_1)\, \mu_1' \leq 1 - \frac{1}{r}\, \theta_1(\mathcal{T})\, \mu_1'(\underline{m}_1)$. In other words, the discount rate is bounded above.

4. We have $\gamma = 1$ on $[l_1, \overline{m}_1]$ (by Lemma 11).

Points 1 and 2 imply that, from any starting point in $[\underline{m}_1, \overline{m}_1]$, m_1 will always reach the centre of the interval $[l_1, \overline{m}_1]$ in finite time. Point 3 implies that the amount of discounting accumulated during this time will be finite. Finally, point 4 implies that the expected present discounted value accumulated from that time on will be strictly positive. ■

Next, we show that the shadow value to coalition 1 of productivity variance in firm 1 is positive.

Lemma 13. $\partial_1^2 S_1 \geq 0$.

Indeed, an increase in variance has two effects in our model, a static effect and a dynamic effect. The static effect is zero, since the flow payoff m_1 from working and the flow payoff k_1 from training are both linear in m_1. The dynamic effect is positive, since μ_1 is convex in m_1.

Proof. The Bellman equation for S_1 is:

$$S_1 = \max_{a_1 \in [0,1]} \left\{ \pi_1^{a_1} + \tfrac{1}{r}\left(\theta_1(a_1)\, \mu_1\, \partial_1 S_1 + \tfrac{1}{2}\, \Psi_1\, \partial_1^2 S_1\right) \right\}.$$

Differentiating twice with respect to m_1 and using the second-order envelope principle, we obtain

$$\begin{aligned}\partial_1^2 S_1 &\geq \partial_1^2 \pi_1^{A_1} + \tfrac{1}{r}\,\theta_1(A_1)\,\left(\mu_1''\,\partial_1 S_1 + 2\,\mu_1'\,\partial_1^2 S_1 + \mu_1\,\partial_1^3 S_1\right) \\ &\quad + \tfrac{1}{2}\,\tfrac{1}{r}\,\left(\Psi_1''\,\partial_1^2 S_1 + 2\,\Psi_1'\,\partial_1^3 S_1 + \Psi_1\,\partial_1^4 S_1\right).\end{aligned}$$

Hence

$$\begin{aligned}\left(1 - \tfrac{1}{r}\,(2\,\theta_1(A_1)\,\mu_1' + \tfrac{1}{2}\,\Psi_1'')\right)\partial_1^2 S_1 &\geq \partial_1^2 \pi_1^{A_1} + \tfrac{1}{r}\,\theta_1(A_1)\,\mu_1''\,\partial_1 S_1 \\ &\quad + \tfrac{1}{r}\,\left(\theta_1(A_1)\,\mu_1 + \Psi_1'\right)\partial_1^3 S_1 + \tfrac{1}{2}\,\tfrac{1}{r}\,\Psi_1\,\partial_1^4 S_1\end{aligned}$$

(on rearranging). That is, $\partial_1^2 S_1$ is a supersolution of the equation

$$\begin{aligned}\left(1 - \tfrac{1}{r}\,\left(2\,\theta_1(A_1)\,\mu_1' + \tfrac{1}{2}\,\Psi_1''\right)\right) X &= \partial_1^2 \pi_1^{A_1} + \tfrac{1}{r}\,\theta_1(A_1)\,\mu_1''\,\partial_1 S_1 \\ &\quad + \tfrac{1}{r}\,\left(\theta_1(A_1)\,\mu_1 + \Psi_1'\right)\partial_1 X + \tfrac{1}{2}\,\tfrac{1}{r}\,\Psi_1\,\partial_1^2 X.\end{aligned}$$

Hence $\partial_1^2 S_1 \geq X$.

Now, for all $\widehat{m}_1 \in [\underline{m}_1, \overline{m}_1]$, we have the representation

$$X(\widehat{m}_1) = \mathrm{E}\left[\int_0^{+\infty} \exp\left(-\int_0^t \beta(m_1(s))\,ds\right)\gamma(m_1(t))\,dt\right],$$

where $m_1(0) = \widehat{m}_1$, $dm_1(t)$ is distributed normally with mean $\zeta(m_1(t))\,dt$ and variance $\eta(m_1(t))\,dt$, and

$$\begin{aligned}\beta &= 1 - \tfrac{1}{r}\,(2\,\theta_1(A_1)\,\mu_1' + \tfrac{1}{2}\,\Psi_1''), \\ \gamma &= \partial_1^2 \pi_1^{A_1} + \tfrac{1}{r}\,\theta_1(A_1)\,\mu_1''\,\partial_1 S_1, \\ \zeta &= \tfrac{1}{r}\,(\theta_1(A_1)\,\mu_1 + \Psi_1'), \\ \eta &= \tfrac{1}{r}\,\Psi_1.\end{aligned}$$

It is easy to see that the discount rate β is bounded. Hence the cumulative discount rate is well defined and finite. Hence the discount factor is well defined, finite and strictly positive. Moreover $\partial_1^2 \pi_1^{A_1} = 0$, $\mu_1'' \geq 0$ (by Property B5) and $\partial_1 S_1 \geq 0$ (by Lemma 7). Hence the flow payoff

$$\gamma = \partial_1^2 \pi_1^{A_1} + \tfrac{1}{r}\,\theta_1(A_1)\,\mu_1''\,\partial_1 S_1 = \tfrac{1}{r}\,\theta_1(A_1)\,\mu_1''\,\partial_1 S_1 \geq 0.$$

Hence the integrand is well defined in the extended sense and positive. (It could in principle be $+\infty$.) Hence $X(\widehat{m}_1) \geq 0$. ■

Next, we show that the shadow value to coalition 1 of productivity variance in firm 1 is finite.

Lemma 14. *$\partial_1^2 S_1 < +\infty$ on $(\underline{m}_1, \overline{m}_1)$.*

Proof. We have

$$\begin{aligned} S_1 &= \max_{a_1 \in [0,1]} \left\{ \pi_1^{a_1} + \tfrac{1}{r} \left(\theta_1(a_1)\, \mu_1\, \partial_1 S_1 + \tfrac{1}{2}\, \Psi_1\, \partial_1^2 S_1 \right) \right\} \\ &= \max_{a_1 \in [0,1]} \left\{ \pi_1^{a_1} + \tfrac{1}{r} \theta_1(a_1)\, \mu_1\, \partial_1 S_1 \right\} + \tfrac{1}{2}\, \tfrac{1}{r}\, \Psi_1\, \partial_1^2 S_1. \end{aligned}$$

Hence

$$\partial_1^2 S_1 = \frac{2\,r}{\Psi_1} \left(S_1 - \max_{a_1 \in [0,1]} \left\{ \pi_1^{a_1} + \tfrac{1}{r} \theta_1(a_1)\, \mu_1\, \partial_1 S_1 \right\} \right)$$

(on rearranging). Now S_1 is bounded (by Lemma 9), $\partial_1 S_1$ is bounded (by Lemmas 7 and 10) and $\Psi_1 > 0$ on $(\underline{m}_1, \overline{m}_1)$. Hence $\partial_1^2 S_1 < +\infty$ on $(\underline{m}_1, \overline{m}_1)$. ■

We are now in a position to prove Lemma 8, which we restate here for the reader's convenience.

Lemma 15. $1 - \frac{1}{r} \big(\widetilde{\theta}_1\, \mu_1 \big)'$ *is bounded above.*

Proof. We have

$$1 - \tfrac{1}{r} \big(\widetilde{\theta}_1\, \mu_1 \big)' = 1 - \tfrac{1}{r}\, \widetilde{\theta}_1\, \mu_1' - \tfrac{1}{r}\, \widetilde{\theta}_1'\, \mu_1.$$

It therefore suffices to show that $-\widetilde{\theta}_1'\, \mu_1$ is bounded above. Let $f_1 : (-\infty, +\infty) \to [0,1]$ be defined by the formula

$$f_1(\lambda) = \operatorname*{argmax}_{a_1 \in [0,1]} \left\{ -\lambda\, a_1 + \theta_1(a_1) \right\}.$$

Then

$$\widetilde{\theta}_1 = \theta_1 \left(f_1 \left(\frac{r\,(m_1 - k_1)}{\mu_1\, \partial_1 S_1} \right) \right)$$

and

$$\widetilde{\theta}_1' = \theta_1'\, f_1'\, r\, \frac{\mu_1\, \partial_1 S_1 - (m_1 - k_1)\, (\mu_1'\, \partial_1 S_1 + \mu_1\, \partial_1^2 S_1)}{(\mu_1\, \partial_1 S_1)^2}.$$

There are now two possibilities. If

$$\theta_1'(\mathcal{T}) \le \frac{r\,(m_1 - k_1)}{\mu_1\, \partial_1 S_1} \le \theta_1'(\mathcal{W}),$$

then we have an interior solution for A_1. In this case,

$$\begin{aligned} -\widetilde{\theta}_1'\, \mu_1 &= -\theta_1'\, f_1'\, r\, \frac{\mu_1\, \partial_1 S_1 - (m_1 - k_1)\, (\mu_1'\, \partial_1 S_1 + \mu_1\, \partial_1^2 S_1)}{\mu_1\, (\partial_1 S_1)^2} \\ &= \frac{-\theta_1'\, f_1'}{\partial_1 S_1} \left(r - \frac{r\,(m_1 - k_1)}{\mu_1\, \partial_1 S_1}\, (\mu_1'\, \partial_1 S_1 + \mu_1\, \partial_1^2 S_1) \right) \\ &\le \frac{-\theta_1'\, f_1'}{\partial_1 S_1} \left(r - \frac{r\,(m_1 - k_1)}{\mu_1\, \partial_1 S_1}\, \mu_1'\, \partial_1 S_1 \right) \end{aligned}$$

(because $f_1' \le 0$ and $\partial_1^2 S_1 \ge 0$)

$$\le \frac{-\theta_1' f_1'}{\partial_1 S_1} \left(r - \theta_1'(\mathcal{W})\, \mu_1' \, \partial_1 S_1\right),$$

(because $\mu_1' \le 0$ and $\frac{r\,(m_1-k_1)}{\mu_1\,\partial_1 S_1} \le \theta_1'(\mathcal{W})$). On the other hand, if

$$\text{either } \frac{r\,(m_1-k_1)}{\mu_1\,\partial_1 S_1} < \theta_1'(\mathcal{T}) \text{ or } \frac{r\,(m_1-k_1)}{\mu_1\,\partial_1 S_1} > \theta_1'(\mathcal{W}),$$

then we have a corner solution for A_1. In this case, $f_1' = 0$ and so $-\widetilde{\theta}_1'\,\mu_1 = 0$. Either way, $-\widetilde{\theta}_1'\,\mu_1$ is bounded above. ■

We conclude with the following complement to Lemma 15.

Lemma 16. $1 - \frac{1}{r}\big(\widetilde{\theta}_1\,\mu_1\big)' > -\infty$ *on* $(\underline{m}_1, \overline{m}_1)$.

Proof. As in the proof of Lemma 15,

$$1 - \tfrac{1}{r}\big(\widetilde{\theta}_1\,\mu_1\big)' = 1 - \tfrac{1}{r}\,\widetilde{\theta}_1\,\mu_1' - \tfrac{1}{r}\,\widetilde{\theta}_1'\,\mu_1.$$

It therefore suffices to show that $-\widetilde{\theta}_1'\,\mu_1$ is finite. If

$$\theta_1'(\mathcal{T}) \le \frac{r\,(m_1-k_1)}{\mu_1\,\partial_1 S_1} \le \theta_1'(\mathcal{W}),$$

then

$$-\widetilde{\theta}_1'\,\mu_1 = -\theta_1'\,f_1'\,r\,\frac{\mu_1\,\partial_1 S_1 - (m_1-k_1)\,(\mu_1'\,\partial_1 S_1 + \mu_1\,\partial_1^2 S_1)}{\mu_1\,(\partial_1 S_1)^2}.$$

But $\partial_1^2 S_1$ is finite on $(\underline{m}_1, \overline{m}_1)$ (by Lemma 14). Hence $-\widetilde{\theta}_1'\,\mu_1$ is finite. If

$$\text{either } \frac{r\,(m_1-k_1)}{\mu_1\,\partial_1 S_1} < \theta_1'(\mathcal{T}) \text{ or } \frac{r\,(m_1-k_1)}{\mu_1\,\partial_1 S_1} > \theta_1'(\mathcal{W}),$$

then $-\widetilde{\theta}_1'\,\mu_1 = 0$. In particular, $-\widetilde{\theta}_1'\,\mu_1$ is finite. ■

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, 539–72.

BECKER, G. S. (1993): *Human Capital*. Third edn. Chicago: The University of Chicago Press.

Bergemann, D., and J. Välimäki (1996): "Learning and Strategic Pricing," *Econometrica*, 64, 1125–1149.

Bolton, P., and C. Harris (1999): "Strategic Experimentation," *Econometrica*, 67, 349–74.

Felli, L., and C. Harris (1996): "Learning, Wage Dynamics and Firm-Specific Human Capital," *Journal of Political Economy*, 104, 838–68.

Heckman, J. (1993): "Assessing Clinton's Program on Job Training, Workfare and Education in the Workplace," Working Paper 4428, N.B.E.R.

Jovanovic, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87, 972–90.

Pissarides, C. A. (1994): "Search Unemployment with On-the-job Search," *Review of Economic Studies*, 61, 457–75.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 70, 2295–2350.

Prendergast, C. (1993): "The Role of Promotion in Inducing Specific Human Capital Acquisition," *Quarterly Journal of Economics*, 108, 523–34.

Stevens, M. (1994a): "A Theoretical Model of On-the-job Training with Imperfect Competition," *Oxford Economic Papers*, 46, 537–62.

——— (1994b): "An Investment Model for the Supply of Training by Employers," *The Economic Journal*, 104, 556–70.

Topel, R. H., and M. P. Ward (1992): "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 107, 439–79.