CHICAGO JOURNALS

The Structure of Wages and Investment in General Training

Daron Acemoglu and Jörn-Steffen Pischke

Massachusetts Institute of Technology

In the human capital model with perfect labor markets, firms never invest in general skills and all costs of general training are borne by workers. When labor market frictions compress the structure of wages, firms may pay for these investments. The distortion in the wage structure turns "technologically" general skills into de facto "specific" skills. Credit market imperfections are neither necessary nor sufficient for firm-sponsored training. Since labor market frictions and institutions shape the wage structure, they may have an important impact on the financing and amount of human capital investments and account for some international differences in training practices.

I. Introduction

The distinction between general and specific skills is the cornerstone of the standard theory of human capital as developed by Becker (1964). Specific skills are useful only with the current employer, whereas general skills are as useful with other employers. In competitive labor markets, workers capture all the returns to their general human capital, and employers have no incentive to pay for investments in these skills. In this paper, we show that if labor market frictions reduce the wages of skilled workers relative to wages of un-

skilled workers (i.e., compress the structure of wages), firms may provide and pay for general training. Credit market problems and the presence of a long-term attachment between the worker and the firm are neither necessary nor sufficient to generate firm-sponsored training. The key is labor market imperfections, which imply that trained workers do not get paid their full marginal product when they change jobs, making *technologically general* skills de facto *specific.*[1]

There is a variety of evidence that suggests that, in line with our approach, firms provide and pay for general training. For example, in Germany, firms voluntarily offer apprenticeships to young workers, and general skills are an important component of these programs as evidenced by the fact that apprentices are given exams by outside boards at the end of their programs. The mere fact that firms provide general training does not establish that they pay the costs since workers may be taking a correspondingly lower wage relative to their marginal product. Nevertheless, most calculations suggest that employers pay for at least part of the costs. For example, the data reported in von Bardeleben, Beicht, and Fehér (1995) show that, even under conservative estimates, the *net* cost of an apprentice to a large German firm is over DM 7,500 a year (see also Harhoff and Kane 1997; Acemoglu and Pischke 1998*b*). Similarly, Ryan (1980) reports sizable net costs of apprenticeship training in a U.S. shipyard. An interesting example of firm-sponsored training is the case of temporary help agencies in the United States, which provide general training to new employees, such as computer and typing skills, and bear the full monetary costs (Krueger 1993; Autor 1998).

The main idea of our paper can be explained using figure 1, which draws the (marginal) product of a worker, $f(\tau)$, as a function of his or her skills, τ. Suppose that workers can quit and work for another firm and, in the process, incur a cost $\Delta \geq 0$. Under the assumption that workers will receive their full product on quitting, their outside option is $v(\tau) = f(\tau) - \Delta$. Suppose that the current employer can keep them by paying this outside option, so their wage is $w(\tau) = f(\tau) - \Delta$. The employer has no incentive to invest in the workers' skills because its profits are $f(\tau) - w(\tau) = \Delta$ irrespective of the value of τ. This is true despite the fact that when $\Delta > 0$, there are mobility

---

[1] In the standard theory, firms pay for skills that are specific, and which skills are specific is determined by technology. In contrast, we focus on skills that are technologically general in the sense that, without frictions, they will be as useful with other employers. Market structure and institutions determine, in equilibrium, which skills are turned into effectively "specific" skills. Becker realized that this may happen when he wrote that "in extreme types of monopsony . . . job alternatives for trained and untrained workers are nil, and all training, no matter what its nature, would be specific to the firm" ([1964] 1993, p. 50), but he did not pursue this further.

f(τ)
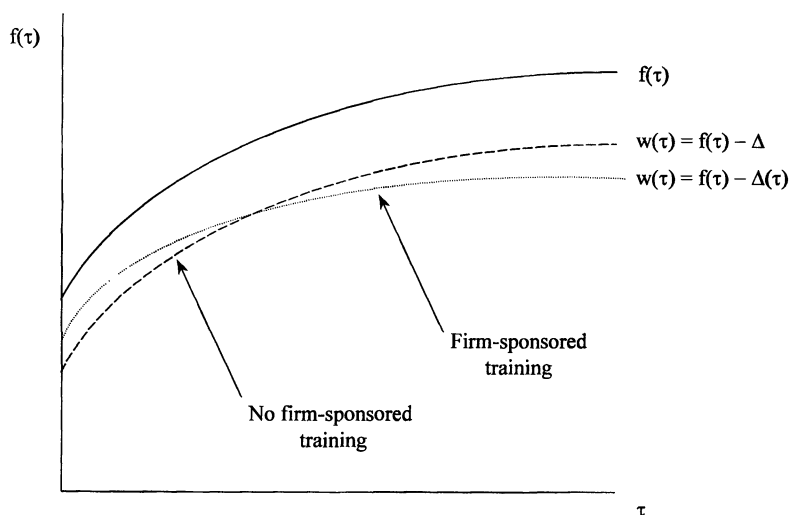
FIG. 1.—Wage structure and training

costs creating an attachment between workers and the firm. Also, notice that a perfectly competitive labor market corresponds to the case with $\Delta = 0$, and employers once again do not invest in general skills. Next, consider the main focus of our analysis, a labor market with a compressed wage structure.[2] Specifically, let the mobility cost be $\Delta(\tau)$, with $\Delta'(\tau) > 0$ as shown with the dotted curve in the figure. Since firms pay workers their outside options, profits are $f(\tau) - w(\tau) = \Delta(\tau)$. Because skilled workers face relatively worse outside opportunities, the equilibrium wage structure is compressed relative to productivity differentials, that is, $w'(\tau) < f'(\tau)$, and the firm makes greater profits from more skilled workers. Therefore, as long as training costs are not too large, the firm will find it profitable to invest in τ. This is the basic story of our paper, which is analyzed in more detail in Section II. In that section, we shall also establish a key comparative static result: even when workers are not credit constrained, as the wage structure becomes more compressed, firms pay for a larger fraction of the costs of general training, and when the structure of wages is sufficiently distorted, they may pay for all the costs.

Our partial equilibrium analysis in Section II assumes that, as in figure 1, the *external wage structure*, $v(\tau)$, is distorted, inducing em-

---

[2] The evidence in Bishop (1987) and Barron, Berger, and Black (1997) shows that wages across workers doing the same job in the same firm differ much less than their productivities, suggesting that the structure of wages in practice is compressed.

ployers to compress the internal wage structure, and shows how this wage compression leads to firm-sponsored training. In Section III, we investigate why a distorted external wage structure may emerge. We show that a range of plausible frictions, such as search, informational asymmetries, and efficiency wages, lead to this type of distortion. Furthermore, even when the labor market is frictionless, complementarities between technologically general and specific skills may induce firms to invest in the general skills of their workers. Finally, we also show that labor market institutions such as union wage setting and minimum wages, which also compress the structure of wages, may encourage firms to invest in the general skills of their employees. Therefore, our model predicts that in a variety of circumstances, we should observe firm-sponsored investments in general training.

The link we draw between labor market institutions and human capital accumulation may be useful in evaluating international patterns in training provision. There are important differences between labor market institutions of Anglo-Saxon economies, continental Europe, and Japan. For example, in contrast to the United States and the United Kingdom, in Germany and Sweden, unions play an important role in wage determination, and there are relatively high wage floors set by minimum wages and unemployment benefits (e.g., OECD 1994). Many economists believe that these institutions compress returns to skills (e.g., Blau and Kahn 1996; Edin and Topel 1997). Comparisons of wage dispersion and returns to education support this view. For example, in the mid-1980s, the log difference of ninetieth and tenth percentile wages was 1.73 in the United States and 1.11 in the United Kingdom as opposed to 0.83 in Germany, 0.67 in Sweden, 1.22 in France, and 1.01 in Japan (OECD 1993). It has also been argued that wage compression reduces not only employment but also investments in human capital (e.g., Lindbeck et al. 1993). In contrast, in our theory a compressed wage structure may induce firms to provide and pay for general training. Therefore, we expect that European and Japanese labor market institutions may increase one of the components of investment in human capital, firm-sponsored general training, and possibly even contribute to total human capital accumulation.[3]

[3] The incidence of company-provided *formal* training appears to be higher in Europe and Japan than in the United States: OECD (1994, table 4.7) reports that 23.6 percent of young workers in France, 71.5 percent of those in Germany, and 67.1 percent of new hires in Japan receive formal training. By way of comparison, only 10.2 percent of U.S. workers receive any formal training during their first 7 years of labor market experience. These data are collected using different methods, however, and are not easily comparable. Using the U.S. National Longitudinal Survey of Youth (NLSY), e.g., Loewenstein and Spletzer (1999) report the incidence of

Since labor market distortions make general skills firm-specific, our model is consistent with a variety of evidence traditionally used to support the presence of firm-specific human capital. But it suggests that many of these apparently specific skills may be more general, which is in line with the fact that in surveys workers claim that most of the training they receive provides skills useful with other employers. For example, Barron, Berger, and Black (1997) find that employers claim that training is valuable with other firms, but productivity growth associated with training exceeds wage growth by a factor of 10. Also, in our economy, as in models of specific training, employers recoup the costs of general training later during the tenure of the workers, which is consistent with Loewenstein and Spletzer's (1998*b*) finding that wage returns to general and specific training provided by the current employer are quite similar. Finally, since training is general, our model predicts an experience premium: wages are higher during the later career of workers because of the investments during the early years. Furthermore, because market frictions make these skills partly specific, there is also a tenure premium (see Altonji and Shakotko 1987; Topel 1991; Altonji and Williams 1997). We discuss other empirical predictions of our analysis and the relevant evidence in Section III.

II. Partial Equilibrium

A. The Environment

We work with the following two-period model throughout this section. In period 1, which we view as the early career of the worker, the employer or the worker or both choose how much to invest in the worker's general human capital, denoted by $\tau \in R_+$. We normalize production during this period to zero and denote the worker's first-period wage by W. In period 2, the worker either stays with the firm at a wage $w(\tau)$ or decides to quit and obtains an outside wage, $v(\tau)$. We also assume that with probability q, the firm and the worker receive an adverse shock, cease to be productive together, and separate. With probability $1 - q$, they can continue their productive relation. Therefore, q is a measure of (expected) turnover in our model. We ignore discounting and assume that all agents are risk-neutral and have preferences defined over the single good of this economy.

Each worker produces output $f(\tau)$ independent of the number

---

formal training as 17 percent. There are also important differences between numbers for formal and informal training. Loewenstein and Spletzer discuss various data sources on informal training for the United States. They find the incidence of informal training to be between 28 percent and 38 percent in the NLSY.

and human capital of other workers.[4] The function $f(\cdot)$ is increasing, differentiable, and concave. If he receives no training, the worker is as productive as during his early career, $f(0) = 0$. The cost of acquiring τ units of skill is $c(\tau)$ in terms of the final good and is incurred by the firm (although the worker can pay for it by having $W < 0$). We assume that $c(\tau)$ is everywhere strictly increasing, differentiable, and convex, with $c'(0) = c(0) = 0$ and $\lim_{\tau \to \infty} c'(\tau) = \infty$. These conditions ensure that the first-best training level, τ^*, is given by $c'(\tau^*) = f'(\tau^*)$, and from the assumptions on the cost of training, $\tau^* > 0$.

The productivity of the worker is the same in all firms since τ is general human capital. The assumption that there are no (technologically) firm-specific skills is extreme but serves to highlight our focus: the presence of frictions may transform technologically general capital into firm-specific human capital. We discuss how technologically specific and general skills interact in Section IIIC.

We distinguish three cases below. In the first, which we call the *constrained regime,* a worker cannot take a wage cut during the first period in order to compensate the firm for the expenses of training, so only the firm makes training investments. The most satisfactory justification for this is contractual problems between the firm and the worker; for example, the employer may not be able to commit to providing training after the worker makes a wage concession. A more common explanation in the literature for why workers do not make contributions to training investments has been the idea that young workers may be credit constrained (see, e.g., Ritzen and Stern 1991). However, since zero output in the first period is only a normalization, it is possible that the worker produces some amount $y_1 > 0$ in the first period and can contribute to training expenses by taking a wage lower than y_1. Nevertheless, such a wage concession would still be costly in the absence of perfect capital markets because it would lead to a nonsmooth consumption profile. Therefore, credit constraints also provide a justification for the constrained regime.

Despite these possible justifications, it should be clear that the constrained regime is an extreme case, and we consider it only to focus on our main innovation—firms' incentives to invest in general training—and to highlight the contrast between perfect and imperfect labor markets. In the second case we consider, the *noncooperative*

---

[4] This assumption is not as restrictive as it appears. For example, if total output is a function of human capital H and physical capital K, $F(H, K)$ exhibits constant returns to scale, and K can be adjusted freely, then the marginal contribution of a worker with human capital τ, $y = f(\tau)$, will be independent of the level of H. The reason is that an optimizing firm will keep the ratio of physical to human capital, K/H, constant.

regime, firms can credibly commit to a level of training, and workers are not credit constrained. However, firms and workers make decisions about training unilaterally, and we allow firms to make positive profits. In particular, we start with a first-period wage $W = 0$, but the worker chooses how much to contribute to training by taking a wage cut. So the level of training is $\tau = c^{-1}(\gamma_f + \gamma_w)$, where c^{-1} is the inverse function of the cost of training, $c(\tau)$; γ_f is the firm's contribution to training costs; and γ_w is the contribution of the worker, so his first-period wage is $W = -\gamma_w$. Our analysis of the noncooperative regime will show that the main results of the constrained regime carry over to this case, demonstrating that contractual problems or credit constraints are not essential for our results.

The third case is the *full-competition regime*, where firms compete in the first period by offering training-wage combinations $\{W, \tau\}$ to workers, and in equilibrium they make zero profits. Contractual problems and credit constraints are once again absent, so $W < 0$ is allowed. This case leads to a number of different results, but our main conclusions that firms may pay for general training and that the extent of their payments depends on the degree of wage compression continue to apply. The assumption that contractual problems are absent in the noncooperative and the full-competition regimes is reasonable when firms have long-term reputations but might be harder to justify in other circumstances, for example when firms are small or face a high probability of failure. We therefore believe that different regimes may be more reasonable descriptions for particular labor markets or episodes.

B. Training in a Perfectly Competitive Labor Market

If a worker quits after the first period, she receives a wage of $v(\tau)$ in the outside labor market. Before discussing labor market frictions, we set the stage by reviewing the case of a competitive labor market. Since all skills are general and the worker can quit at no cost, we have $v(\tau) = f(\tau)$. This implies that workers have to be paid their full marginal product,[5] that is, $w(\tau) = v(\tau) = f(\tau)$. The following result is immediate.

PROPOSITION 1. Suppose that we are in the constrained regime and labor markets are competitive. Then the equilibrium training level is $\hat{\tau} = 0$ (and $W = 0$).

In this case, the worker cannot contribute to training costs, so the firm chooses the level of training unilaterally. Since $w(\tau) = f(\tau)$, the firm cannot recoup the costs of training during the later career of

---

[5] Recall that the marginal product of the worker is equal to $f(\tau)$, not $f'(\tau)$.

the worker, so no investment takes place, even though the optimal amount of training, τ^*, is strictly positive. It is sometimes asserted that credit constraints faced by workers may induce firms to invest in general training. Proposition 1 shows that such constraints are not sufficient for firm-sponsored training.

We next analyze the noncooperative and full-competition regimes.

PROPOSITION 2. In both the noncooperative and the full-competition regimes, when labor markets are competitive, the equilibrium level of training is $\hat{\tau} = \tau^*$ and $W = -c(\tau^*)$.

Consider first the noncooperative regime. Here the worker realizes that the firm will not contribute to training, so $\gamma_f = 0$ and $\tau = c^{-1}(-W)$. Hence, the worker chooses $W = -c(\tau^*)$, which induces training τ^*, and achieves the highest lifetime payoff. Intuitively, since the worker is the full residual claimant of returns from training, he has the right incentives to invest.

In the full-competition regime, firms compete to attract workers in the first period and are forced to offer the highest lifetime utility, which in this case is $f(\tau^*) - c(\tau^*)$. So firms provide training τ^* and ensure that the worker pays for the cost by offering $\{W, \tau\} = \{-c(\tau^*), \tau^*\}$. Workers, who are paying the costs of general training, again have the right incentives to invest, and the first-best level of training is achieved.

Therefore, when labor markets are competitive and workers are allowed to contribute to training, the equilibrium achieves first-best training and the worker bears the full costs, as emphasized in Becker's (1964) seminal analysis. Note that the presence of separations with probability q is of no consequence because the worker gets exactly the same returns for his general human capital in the outside market.

C. Frictional Labor Markets in the Constrained Regime

We now model frictional labor markets by assuming that $v(\tau) < f(\tau)$. Despite the fact that τ is general human capital, when the worker separates from his employer, he receives a wage lower than his marginal product. In the next section, we discuss in detail how different types of frictions and institutions determine $v(\tau)$ and its relation to $f(\tau)$. For now, we take $v(\tau)$ as given and assume that $v''(\tau) \geq f''(\tau)$, which is a sufficient (but not necessary) restriction for the second-order conditions to hold. In the specific examples in Section III, this restriction will hold. Since $v(\tau) < f(\tau)$, there is a surplus that the firm and the worker can share when they are together. For the exposition in this section, we adopt the Nash bargaining approach. We

also start with the constrained regime and return to the other cases below.

Asymmetric Nash bargaining and risk neutrality imply that $w(\tau)$, the second-period wage at the current firm, is

$$w(\tau) = v(\tau) + \beta[f(\tau) - v(\tau) - \pi_0], \qquad (1)$$

where $\beta \in [0, 1]$ is the bargaining power of the worker, and π_0 is the outside option of the firm, which we normalize to zero. The equilibrium wage rate $w(\tau)$ is independent of the cost of training, $c(\tau)$. This is a feature of the temporal structure of our economy. The level of training is chosen by the firm, and then the worker and the firm bargain over the wage rate. At this point, training costs are already sunk.

Profits of the firm are

$$\pi(\tau) = (1 - q)[f(\tau) - w(\tau)] - c(\tau)$$
$$= (1 - \beta)(1 - q)[f(\tau) - v(\tau)] - c(\tau),$$

where we have incorporated the fact that, with probability q, there will be an involuntary separation. In this regime the firm decides the level of training and bears all the costs, so it chooses τ to maximize $\pi(\tau)$, which gives the first-order condition

$$(1 - \beta)(1 - q)[f'(\hat{\tau}) - v'(\hat{\tau})] - c'(\hat{\tau}) = 0. \qquad (2)$$

The necessary condition for the firm to invest in the general human capital of the worker, that is, for $\hat{\tau} > 0$, is $\pi'(0) > 0$. Since $c'(0) = 0$, firms will invest in training if and only if $f'(0) > v'(0)$ and $(1 - \beta)(1 - q) > 0$.

PROPOSITION 3. Suppose that we are in the constrained regime, labor markets are frictional, $\beta < 1$, and $q < 1$. Then as long as $f'(0) > v'(0)$, the firm invests a positive amount in general skills, that is, $\hat{\tau} > 0$.

In contrast to the case of competitive labor markets, the firm may now have an incentive to invest in the general skills of its workers. The condition $f'(0) > v'(0)$ implies that the wage structure is compressed (at the point of $\tau = 0$), so an increase in the worker's productivity increases profits, encouraging the firm to invest in training.[6] What is relevant to the firm is the wage it pays, $w(\tau)$, that is, the internal wage structure. However, the internal wage structure is endogenous and is linked to the external wage structure, $v(\tau)$. In particular, the wage rule (1) implies $w'(\tau) = \beta f'(\tau) + (1 - \beta)v'(\tau)$.

---

[6] The additional requirements that $\beta < 1$ and $q < 1$ ensure that the firm gets some rents from the relation and that the employment relationship does not end with probability one.

Therefore, $f'(\tau) > v'(\tau)$ is equivalent to $f'(\tau) > w'(\tau)$, so that wages increase less with skills than productivity does, and the firm makes higher profits from trained workers. In other words, the internal wage structure is distorted only when the external wage structure is.[7] Note, however, that $v'(\tau) < f'(\tau)$ does *not* imply that training is less productive with other firms: since τ is general skills, the worker produces $f(\tau)$ with outside firms; but moving to a new firm is costly, and more so for more skilled workers. More explicit microfoundations for these costs will be given in the next section.

Although a distorted wage structure encourages firms to pay for training, equilibrium training, $\hat{\tau}$, is generally less than the first-best, τ^*. In particular, as long as $\beta > 0$ and $v'(\tau^*) > 0$, or if $q > 0$, equation (2) implies that $\hat{\tau} < \tau^*$.

A key comparative static result is immediate from our analysis so far. Let $v(\tau) = a\tilde{v}(\tau)$. Then everything else being equal, a reduction in a increases firms' investments in training, $\hat{\tau}$ (see eq. [2]). A decrease in a reduces the outside option of skilled workers relative to the outside opportunities of the unskilled, compressing the wage structure. This implies that the firm can capture additional rents from the skilled, so it invests more in its employees' skills. Therefore, contrary to conventional wisdom, a more compressed wage structure may improve human capital investments.

This result implies that the distortion of the wage structure may actually improve welfare. This is the well-known theory of the second-best at work. Since in the constrained regime training outcomes are inefficient, another distortion, in this case in the labor market, may induce firms to undertake some of these investments and improve output and welfare. For example, a move from $v(\tau) = f(\tau)$ to $v(\tau) = af(\tau) + b$ with $a < 1$ increases human capital investments and does not affect other margins, so it increases net output (since $\hat{\tau} = 0 < \tau^*$). Naturally, in practice, increased frictions will have a number of allocative costs, such as lower employment. These costs need to be compared to the benefits in terms of better training incentives. Also our simple example in which workers cannot take wage cuts to bear the costs of training exaggerates the potential benefits from a distortion in the wage structure (see the next two subsections). In any case, the implications of labor market frictions on

[7] This is a feature of Nash bargaining. Other bargaining solutions give similar results but make the dependence of the internal on the external wage structure less transparent. Notice also that $v(\tau)$ and $w(\tau)$ are the wage structures "off the equilibrium path" because all workers and firms are homogeneous. So they invest the same amount, and in equilibrium we observe only $v(\hat{\tau})$ and $w(\hat{\tau})$. It is straightforward, but not very instructive, to introduce worker or firm heterogeneity so that in equilibrium we observe different workers paid different wages.

training are worth bearing in mind when suggesting labor market reforms. For example, proposals for reducing union power and removing other regulations in the German labor market, which are on the current political agenda, could have unforeseen consequences regarding the German apprenticeship system, where employers pay for the general training of their workers.

Another useful comparative static result pertains to turnover, q. Equation (2) immediately implies that $d\hat{\tau}/dq < 0$, so turnover reduces training. The reason is that the firm benefits from training only when the worker does not change jobs, and higher turnover makes this less likely. Since the equilibrium level of training, $\hat{\tau}$, is already less than the first-best τ^*, an increase in q makes training suboptimally low. It is often argued that high-turnover economies such as the United States do not generate sufficient investments in worker skills and that this represents an important market failure (e.g., Blinder and Krueger 1996). Indeed, cross-sectional comparisons reveal that high-turnover countries or industries have lower training. For example, Topel and Ward (1992) find that the median number of jobs held by a male worker with 10 years of experience is six in the U.S. labor market, whereas it is one (Acemoglu and Pischke 1998b) or two (Dustmann and Meghir 1997) in Germany, where young workers are much more likely to receive formal training (see also OECD 1994). Our model explains these correlations and suggests why high turnover causes less training in general skills, and why this may represent a market failure. Standard theory predicts a negative correlation between specific skills and turnover but suggests that such a negative correlation is optimal.

While we find a link between general training and turnover, it should be stressed that it is not the attachment between firms and workers that leads to firm-sponsored training. To see this, suppose that $q = 0$ and outside employers offer wages equal to $f(\tau)$, but there is a cost of moving to a new employer, Δ, so that the worker receives $v(\tau) = f(\tau) - \Delta$. In this case, $w(\tau) = f(\tau) - (1 - \beta)\Delta$, and all workers stay with their initial firms. Although workers never leave their employer, there is no firm-sponsored investment in training because there is no distortion in the wage structure, that is, $v'(\tau) = w'(\tau) = f'(\tau)$.

It is also worth noting that when $v'(0) < f'(0)$ so that firms invest in training, there is both an experience premium and a tenure premium. The experience premium, conditional on tenure, is given by the change in wages for a worker who switches employers, that is, $EP = v(\hat{\tau}) - W$. Since $W = 0$, $EP > 0$ except in the extreme case in which the outside wage structure does not reflect any of the general skills. The tenure premium, on the other hand, is the additional

wage increase that workers staying with their initial employers receive compared to switchers, which in this case is equal to $w(\hat{\tau}) - v(\hat{\tau}) = \beta[f(\hat{\tau}) - v(\hat{\tau})]$. Estimates in the literature suggest that an increase in profit per worker increases wages, and the coefficient, which corresponds to β, varies between 0.003 and 0.3 (see Abowd and Lemieux 1993; Blanchflower, Oswald, and Sanfey 1996). This suggests possible tenure effects ranging from quite small to sizable, consistent with empirical evidence that finds different tenure effects depending on specification and sample (e.g., Altonji and Shakotko 1987; Topel 1991; Altonji and Williams 1997).

   Proposition 3 shows that firms prefer to pay for the training rather than employ an unskilled work force when wage differentials are compressed. However, there might be another, more profitable, strategy, which is to hire (poach) trained workers in the second period. In order to understand this poaching problem, notice that workers have no incentive to quit in equilibrium since $w(\tau) > v(\tau)$, but firms would like to hire trained workers because $w(\tau) < f(\tau)$. Whether poaching trained workers from other firms is profitable or not depends on the source of the distortion causing $v(\tau) < f(\tau)$. Since we take the external wage structure as given, we delay a more detailed discussion of this possibility until we analyze more specific mechanisms in the next section.

D.   Firm-Sponsored Training in the Noncooperative Regime

We now discuss the impact of labor market frictions on training when both firms and workers can contribute to training investments. We find that, contrary to common beliefs, credit market problems are *not* necessary for firms to bear the cost of general training. Whether they do or not is once again determined by the structure of wages.

   Recall that in this regime, training investments by firms and workers are chosen *noncooperatively*. In particular, we start with a first-period wage $W = 0$, and the worker and the firm simultaneously choose the amount of money they wish to spend on training, γ_w and γ_f. The amount of training is τ_{nc} such that $c(\tau_{nc}) = \gamma_w + \gamma_f$ or $\tau_{nc} = c^{-1}(\gamma_w + \gamma_f)$, and the worker's first-period wage is $W = -\gamma_w$. Therefore, the worker maximizes $v(\tau_{nc}) + (1 - q)\beta[f(\tau_{nc}) - v(\tau_{nc})] - \gamma_w$ by choosing $\gamma_w \geq 0$ and takes γ_f as given. Intuitively, with probability $1 - q$, the worker stays with the firm at the wage $w(\tau) = \beta f(\tau) + (1 - \beta)v(\tau)$. With probability q, he is forced to quit and receives $v(\tau)$. The first-order condition for the worker's contribution is

$$v'(\tau_{nc}) + (1 - q)\beta[f'(\tau_{nc}) - v'(\tau_{nc})] - c'(\tau_{nc}) = 0 \quad \text{if } \gamma_w > 0$$
$$\leq 0 \quad \text{if } \gamma_w = 0. \tag{3}$$

Similarly, the firm maximizes $(1 - q)(1 - \beta)[f(\tau_{nc}) - v(\tau_{nc})] - \gamma_f$ by choosing $\gamma_f \geq 0$ and taking γ_w as given. The first-order condition for the firm is

$$(1 - q)(1 - \beta)[f'(\tau_{nc}) - v'(\tau_{nc})] - c'(\tau_{nc}) = 0 \quad \text{if } \gamma_f > 0$$
$$\leq 0 \quad \text{if } \gamma_f = 0, \tag{4}$$

which is essentially the same as (2). Inspection of equations (3) and (4) implies that, generically, only one of them holds as an equality, so one of the parties bears the full cost of training. The reason is that the contributions of the worker and the firm are perfect substitutes. More precisely, let τ_w be the level of training that satisfies (3) as an equality and τ_f be the solution to (4) as an equality. Then we get the following proposition.

PROPOSITION 4. Suppose that we are in the noncooperative regime. If $\tau_f > \tau_w$, then the firm bears all the cost of training, $W = \gamma_w = 0$, and $\tau_{nc} = \tau_f$. In contrast, if $\tau_w > \tau_f$, then $\gamma_f = 0$, the worker bears all the cost of training, $W = \gamma_w = -c(\tau_{nc})$, and $\tau_{nc} = \tau_w$.

Despite the fact that training is general and the worker is *not* credit constrained, the firm may bear all the costs of training. When $\tau_f > \tau_w$, the results in Section IIC continue to hold. Therefore, for our results that firms pay for general training to be true, we do not need the assumption that workers cannot contribute to the costs of training: as long as the structure of wages is sufficiently distorted, firms will be more willing to invest in training than workers, and they will bear all the costs.

An important result of this analysis is that the more distorted the wage structure is (i.e., the lower v' is relative to f'), the more likely the firm, rather than the worker, is to pay for training. Therefore, our model predicts that in economies with compressed wage structures such as Germany and Sweden, employers should pay for general training, whereas in the United States it may be the workers who bear the cost of a range of training investments (such as vocational courses). Also, when the firm is paying for training, a further distortion in the wage structure increases training, whereas when workers are bearing the costs, a distortion in the structure of wages reduces training. Finally, inspection of (3) and (4) shows that a larger bargaining power for the firm, that is, a lower value of β, makes it more likely that the firm will finance the costs of training. The reason is that for a given $v(\tau)$, a decline in β makes the *internal wage structure*

more compressed, encouraging the firm to make a larger contribution to training.

Hashimoto (1979, 1981) and Hashimoto and Yu (1980) have previously analyzed how costs of investments in specific training are shared between the firm and the worker. Since labor market imperfections turn τ into de facto specific skills, our analysis is related to this work. Hashimoto's work assumes that there are transaction costs, causing inefficient separations later in the career of workers. In particular, workers and firms receive idiosyncratic taste and productivity shocks and can unilaterally decide to end the relation. The firm-worker pair therefore shares productivity gains to minimize inefficient separations, and this sharing rule for returns from training dictates how the costs of investments should be shared. Our analysis differs in that there are no transaction costs after investment and, thus, no inefficient separations in the second period. However, one might view our model as including "transaction costs" at the point of training, which force workers and firms to make their contributions to training investments *noncooperatively* (in the next subsection, we remove these transaction costs). This feature implies that the expectation of future returns, partially shaped by outside wages, determines the willingness of the parties to invest. When the external wage structure is highly distorted, the resulting internal wage structure does not reward the worker for his skills, so the firm has to pay all the cost. In this setup, either the firm or the worker pays for all the costs because investments by the two parties are perfect substitutes. With a modified setup in which $\tau = h(\gamma_w, \gamma_f)$ and the cross-partial derivative of h is nonzero, both the firm and the worker contribute to training as in Hashimoto's papers.

E.   Training Investments in the Full-Competition Regime

Now workers can contribute to training costs and firms can commit to providing training, but in contrast to the noncooperative regime, firms compete for workers by offering wage-training packages, $\{W, \tau\}$, taking workers' valuation of the training into account. Firms maximize profits, which are (by substitution for the wage rule [1])

$$\pi(\tau, W) = (1 - q)(1 - \beta)[f(\tau) - v(\tau)] - c(\tau) - W, \quad (5)$$

by choosing W and τ, subject to the constraint that workers receive as much utility as that offered by other firms, U, which is

$$v(\tau) + (1 - q)\beta[f(\tau) - v(\tau)] + W \geq U. \quad (6)$$

Competition ensures that U is high enough so that $\pi = 0$.[8] This implies the following proposition.

PROPOSITION 5. In the full-competition regime, all firms offer first-period wage-training combination $\{W_{fc}, \tau_{fc}\}$ such that $c'(\tau_{fc}) = (1-q)f'(\tau_{fc}) + qv'(\tau_{fc})$ and $W_{fc} = (1-q)(1-\beta)[f(\tau_{fc}) - v(\tau_{fc})] - c(\tau_{fc})$.

To understand the level of training in this case, note that full competition induces a "cooperative" choice of training.[9] In particular, in the noncooperative regime, the firm considered only its own profits when choosing its contribution to training. However now, via the participation constraint of the worker, (6), the firm takes into account how much the utility of the worker increases as a result of a marginal increase in training. This turns the problem into one of maximizing joint surplus, $(1-q)f(\tau) + qv(\tau)$ (i.e., with probability $1 - q$, the pair remains together and the joint return is $f(\tau)$; and with probability q, there is a separation, and the firm receives no return from training whereas the worker receives $v(\tau)$).

Notice that as long as $q > 0$, the external wage structure continues to matter for training. However, a more distorted external wage structure now reduces training. Also when $q > 0$ and $f'(\tau) > v'(\tau)$, the level of training investment is generally less than the first-best amount, because with probability q the worker is employed by another firm. So ex ante investments create positive externalities on his potential future employers (this is not the case when $f'(\tau) > v'(\tau)$ is caused by complementarities between general and specific skills as in Sec. III C but is true in the other cases discussed in the next section). In particular, if $v(\tau)$ is the wage that future employers pay the worker, then their profit is equal to $f(\tau) - v(\tau)$, and a higher level of τ increases future employers' profitability. The worker and the firm do not take this into account in their training decisions, making training suboptimally low (see Acemoglu [1997] for further details).

More important, notice that as long as β is sufficiently less than one, we have $W > 0$. Therefore, the worker receives a positive salary, despite the fact that he is not credit constrained and his net output is $-c(\tau_{fc})$. In particular, the larger the gap between marginal product and the outside wage at the level of equilibrium training, $f(\tau_{fc}) - v(\tau_{fc})$, the greater the first-period wage. The gap $f(\tau_{fc}) - v(\tau_{fc})$ tends

---

[8] Otherwise, another firm would offer $U + \epsilon$, for ϵ sufficiently small and positive, attract all the workers, and make positive profits.

[9] We obtain very similar results when W and τ are determined by Nash bargaining in the first period. In particular, the presence of the first-period wage W makes utility fully transferable, so the choice of training would be "cooperative," i.e., maximize joint surplus.

to be larger when the external wage structure is more distorted (especially if we start with $f(0) = v(0)$). Therefore, our analysis with full competition also suggests that when the wage structure is more distorted, firms pay for a larger fraction of training costs. Intuitively, with imperfect labor markets, the worker has an attachment to his employer, enabling the firm to make positive profits in the second period. Competition then forces the firm to pay these profits in the first period. The firm's monopsony power in the second period originates from the gap between $f(\tau_{fc})$ and $v(\tau_{fc})$. So the more distorted $v(\tau)$ is (or the lower $v(\tau_{fc})$ is), the greater the rent the firm expects, and competition forces it to pay a larger fraction of the costs of training. In fact, if $f(\tau_{fc}) = v(\tau_{fc})$ as in competitive markets, the firm earns no rents in the second period, so $W = -c(\tau_{fc})$; that is, the firm pays nothing toward general training. Therefore, the important conclusion overall is that although the assumption of full competition modifies our analysis, as long as the labor market is imperfect, firms continue to contribute to the cost of general training, and their contribution tends to be larger when the (external) wage structure is more compressed. Hence, except for the impact of wage compression on the total amount of training, our qualitative results from the constrained regime continue to hold.[10]

Our analysis of full competition is closer in spirit to Hashimoto's work. In particular, now there are no transaction costs in the first period, and so the level of training is chosen "cooperatively." While there are no inefficient separations, the firm cannot choose and commit to future wages, so returns from training are shared by Nash bargaining in the second period. Given this sharing rule, training costs are shared in the first period so as to ensure zero profits for the firm.

III. Specific Mechanisms and the Role of Institutions

The previous section described our simple theory of firm-sponsored investment in general training. The key ingredient was a compressed wage structure such that $f'(\tau) > w'(\tau)$. We found that the crucial condition to ensure this is $f'(\tau) > v'(\tau)$; that is, outside opportunities for the worker should improve less than his productivity as he acquires more skills. Although the structure of wages is taken as given by the firm and the worker, it is an equilibrium object. In this

---

[10] Also notice that the tenure premium, $\beta[f(\tau_{fc}) - v(\tau_{fc})]$, is again positive. The experience premium, $v(\tau_{fc}) - W_{fc}$, on the other hand, may be negative, but if β is sufficiently high, it will be positive.

section we discuss how a range of plausible labor market frictions lead to a distortion in the external wage structure, $v(\tau)$, inducing firm-sponsored training. Throughout this section, we simplify our discussion by focusing on the constrained regime, so we study only firms' incentives to invest in skills. Our analysis in Sections IID and IIE shows that firms continue to pay for general training in the noncooperative and full-competition regimes. But in the full-competition regime a distortion in the external wage structure reduces training, whereas in the noncooperative regime it may reduce or increase investments. This has to be borne in mind when one interprets our results in this section. Our aim here is to bring out the major ideas rather than analyze each model fully. For this reason we keep the exposition as simple as possible.

A.   Search and Monopsony

Consider the same setup as in Section II, but in the second period the worker has to find a new firm if he quits. With probability p_w, the worker is successful and finds a new employer; with probability $1 - p_w$, he is unemployed and receives unemployment benefit $b(\tau)$. If he finds an employer, he has to bargain with this firm to determine wages. The worker's outside option in this second and final bargain is zero. With the same bargaining power, β, for the worker as above, he will get a wage $w_2(\tau) = \beta f(\tau)$, and his new employer will capture a proportion $1 - \beta$ of the output. The fact that there is no further period is a special, but nonessential, feature. In the Appendix we analyze the infinite-horizon case and establish the same results.

   The outside option of the worker in the bargain of the first period is therefore $v(\tau) = p_w \beta f(\tau) + (1 - p_w) b(\tau)$. The first-order condition for the firm's investment in training is therefore $(1 - \beta)[(1 - p_w\beta)f'(\tau) - (1 - p_w)b'(\tau)] = c'(\tau)$. As in Section II, firms invest in general training if the external wage structure is distorted against skilled workers (i.e., $f'(0) > v'(0)$). In this model, this is equivalent to $p_w\beta f'(0) + (1 - p_w)b'(0) < f'(0)$, which will be satisfied if $b'(0) < f'(0)$ and $p_w < 1$, or $\beta < 1$. Most unemployment insurance systems are progressive, so $b'(0) < f'(0)$ is a weak requirement, and $p_w < 1$ and $\beta < 1$ are almost always true in models with frictions. Therefore, under fairly weak conditions, this model predicts firm-sponsored investments in general training.

   We refer to this situation as search-induced monopsony: because it is costly for the worker to change employers, the firm has some monopsony power and captures part of the higher output due to the worker's higher productivity. There are two costs of leaving the current employer that underlie the monopsony power of the firm.

First, the worker anticipates that his future employers will capture a certain fraction of his productivity, so the monopsony power of potential future employers contributes to the monopsony power of the current employer. Second, a worker who quits can suffer unemployment, which reduces the return to quitting. Both costs increase with skills, which compresses the equilibrium wage structure and induces firm-sponsored training (see the Appendix and also Acemoglu [1997] for further details).

This model predicts that when the labor market is more "frictional" in the sense that p_w, the exit rate from unemployment, is lower and unemployment higher, we should observe more firm-sponsored training. The reason is that a lower exit rate from unemployment, p_w, makes the wage structure more distorted since $b'(\tau) < f'(\tau)$. According to OECD (1993), monthly exit rates from unemployment are 48.2 percent in the United States, 22 percent in Japan, 7.6 percent in Germany, and 6.7 percent in France. Therefore, in line with our theory, the numbers reported in note 3 suggest that economies with more frictional markets may have more firm-sponsored formal training programs.

It is also instructive to contemplate whether poaching of trained workers may change the implications of this specific mechanism. The answer depends on the exact specification of the poaching process. A reasonable first pass is to consider a situation in which a new firm contacts and makes a poaching offer to an already-employed trained worker. This will create Bertrand competition between the two firms for this particular worker, increasing the worker's wage to his marginal product. Recall that in this model there are search frictions, so it is plausible that it would be costly for firms to find trained workers to make poaching offers to. Hence, anticipating that after a poaching offer there are no profit opportunities, firms would not make such offers. Therefore, introducing the possibility of poaching in the simplest way does not affect our main results.

B.   Asymmetric Information

Skills may be technically general, but outside employers may be unable to ascertain whether a worker actually possesses these skills, or in what amount or quality. If this is the case, the outside wage will not reflect these uncredentialed skills, or not reflect them fully so that $f'(\tau) > v'(\tau)$. This has been suggested by Katz and Ziderman (1990) and analyzed by Chang and Wang (1996). Bishop (1994) finds empirical support for this notion using data from the National Federation of Independent Business Survey.

Information advantages of the incumbent employers may lead to

firm-sponsored training even if the skills are observable. For example, the content of German apprenticeship programs is well known; thus τ is observed by outside firms, but the initial employer still has superior information regarding the ability of its workers. We have analyzed this case in Acemoglu and Pischke (1998b). The following adverse selection model is based on our previous work.

Workers have two different abilities denoted by η. A proportion p have low ability normalized to $\eta = 0$. The remaining proportion $1 - p$ have high ability with $\eta = 1$. The production function is $f(\tau, \eta) = \tau\eta$. The incumbent firm does not know the ability of a particular worker at the beginning of period 1 when it must decide about training. At the end of period 1, it learns the worker's type and offers a wage that can be contingent on ability, $w(\tau, \eta)$. Outside firms do not know worker ability but observe the level of training the worker has received. They offer a wage, $v(\tau)$, conditional on training. Workers quit their original employer whenever the outside wage is higher, that is, when $v(\tau) > w(\tau, \eta)$. We also assume that there are other (exogenous) reasons for quits, so that even when $w(\tau, \eta) \geq v(\tau)$, workers separate with probability λ.

To avoid issues of bargaining with asymmetric information, we give all the bargaining power to the incumbent firm by setting $\beta = 0$. Therefore, the firm will offer a wage $w(\tau, \eta = 0) = 0$ to low-ability workers and the lowest possible wage to high-ability workers that will prevent them from leaving, that is, $w(\tau, \eta = 1) = v(\tau)$. At this wage, only the fraction λ of the high-ability workers who are unhappy in this firm would quit. The outside market is competitive but, as noted above, cannot distinguish high-ability workers, so the outside wage is equal to the expected productivity of workers who separate. Since some high-ability workers quit (i.e., $\lambda > 0$), we have $v(\tau) > 0$. This implies that all low-ability workers will also quit to take advantage of the higher outside wage. In equilibrium, expected productivity and the wage in the outside market are

$$v(\tau) = \frac{\lambda(1 - p)\tau}{p + \lambda(1 - p)}.$$

The incumbent employer keeps a fraction $(1 - \lambda)(1 - p)$ of workers, all of whom are high-ability. Therefore, profits are given by

$$\pi(\tau) = (1 - \lambda)(1 - p)[\tau - w(\tau, 1)] - c(\tau)$$
$$= (1 - \lambda)(1 - p)[\tau - v(\tau)] - c(\tau).$$

In words, the firm pays the cost of training for all workers because worker ability is not observed before training. After training, all low-ability and a proportion λ of high-ability workers leave and the firm

pays $v(\tau)$ to the remaining workers and makes profits equal to $\tau - v(\tau)$ per retained worker. Therefore, the first-order condition for training is

$$\pi'(\tau) = (1 - \lambda)(1 - p)[1 - v'(\tau)] - c'(\tau) = 0. \qquad (7)$$

The firm retains only highly skilled workers, so $f'(\tau) = 1$. Since we also have $c'(0) = 0$, the necessary and sufficient condition for firm-sponsored training is $v'(0) < 1$, our familiar condition that the wage structure should be compressed. It is immediate to see that this condition is always satisfied because $v'(\tau) = \lambda(1 - p)/[p + \lambda(1 - p)] < 1$. Intuitively, the presence of low-ability workers in the second-hand labor market implies that firms view workers in this market as "lemons." They are therefore unwilling to increase their wage offers by much for workers with higher τ because training is not useful to low-ability workers, who are the majority of those in the secondhand market.

Many of the assumptions in this example are inessential and were made only to simplify the exposition. The crucial ingredient is that training and ability are complements, as captured by the multiplicative production function $f(\tau, \eta) = \tau\eta$. To see the importance of complementarity between unobserved ability and training, consider instead $f(\tau, \eta) = \tau + \eta$. The outside wage in this case is

$$v(\tau) = \frac{p\tau + \lambda(1 - p)(1 + \tau)}{p + \lambda(1 - p)} = \tau + \frac{\lambda(1 - p)}{p + \lambda(1 - p)}.$$

The outside wage now increases one for one with τ, that is, $v'(\tau) = 1$. Therefore, (7) is satisfied at $\tau = 0$, and the firm does not invest in the training of its workers. The reason is that training raises the productivity of the more and less able workers by an equal amount. Asymmetric information still leads to rents for the incumbent firm, but it does *not* lead to a distortion of the wage structure.

In Acemoglu and Pischke (1998b), we discuss the case in which outside firms can make poaching offers, and we show that the results discussed here are robust to this extension. The intuition is that the superior information of the incumbent firm creates a "winner's curse": the incumbent would stop competing against the raiding firm only when the raider's offer exceeds the worker's productivity. As a result, the raider will attract the worker only when his productivity falls short of his wage. Also in that paper we present empirical evidence for adverse selection among German apprentices. We show that apprentices who leave their training firm because of the military draft (an exogenous separation) earn more than those who stay at the apprenticeship firm and other quitters. Unlike other quitters

and stayers, military quitters are freed from the adverse selection problem because the reason for their separation is observed by the outside market.

C. Firm-Specific Human Capital

Our analysis has so far concentrated on general human capital for clarity. However, it is undoubtedly true that there exist skills that are more useful in the current firm than in outside firms. Becker's (1964) classic analysis discussed investment in such skills and concluded that the firm should pay for at least part of the costs, and Hashimoto (1979, 1981) showed how the costs of these investments are shared. In the presence of purely specific skills, markets are not competitive in the most usual sense; if a worker has some skills that can be used in only one firm, then for one of the commodities there is only one buyer and one seller, so price-taking behavior does not apply. In this subsection we show that this deviation from pure competition also leads to firm-sponsored investments in general training.

Assume that output in the second period is now given by $y = f(\tau, s)$, where s is firm-specific human capital. Once again, in the first period the firm chooses τ at cost $c(\tau)$, and there is an exogenous probability q that the pair will separate in the second period. The source of the firm-specific skill, s, is inessential: it could be acquired during the first period that the worker spends with the firm (e.g., via a learning mechanism as in Jovanovic [1979]), so $s > 0$. Alternatively, the firm may choose how much to invest in these skills with some cost function $\phi(s)$ such that $(1 - q)\partial f(\tau, 0)/\partial s > \phi'(0)$, which ensures that the firm would always like to invest a positive amount in firm-specific skills. Both scenarios are equivalent for the purposes of this section.

Since there is competition among outside firms, we have $v(\tau) = f(\tau, 0)$. Generally, $v(\tau)$ is independent of s because s is useful only in the current firm. This is the crucial ingredient creating a distortion in the internal wage structure. Assuming Nash bargaining once more, we have $w(\tau, s) = \beta f(\tau, s) + (1 - \beta)f(\tau, 0)$. The firm will solve

$$\max_{\tau} \pi(\tau, s) = (1 - q)\,[\,f(\tau, s) - w(\tau, s)\,] - c(\tau)$$
$$= (1 - q)\,(1 - \beta)\,[\,f(\tau, s) - f(\tau, 0)\,] - c(\tau).$$

As before, this implies that the firm invests $\hat{\tau} > 0$ only if $\beta < 1$, $q < 1$, and $\partial f(0, s)/\partial \tau > v'(0)$ or if $\partial f(0, s)/\partial \tau > \partial f(0, 0)/\partial \tau$. Therefore, for firm-sponsored investment in general training, we need

$\partial^2 f(\tau, s)/\partial \tau \partial s > 0$, that is, a *complementarity between firm-specific and general skills*. In fact, since $c'(0) = 0$, it is necessary and sufficient for firm-sponsored investments in general training that $\partial^2 f(\hat{\tau}, s)/\partial \tau \partial s > 0$, $q < 1$, and $\beta < 1$.

To summarize, if firm-specific skills and general skills are complements in the production function, increasing general skills raises productivity more than outside wages, encouraging the firm to invest in these general skills.[11] If specific and general skills do not interact, the outside wage function has the same slope in τ as the production function. In this case, specific skills generate rents from the current employment relationship, but these rents are the same at all levels of skill. The firm therefore has no incentive to invest in general skills.

Notice also that standard theory suggests that there should be less investment in firm-specific skills when turnover is higher. Our model points out that there should be more firm-sponsored investments in general skills when there are more firm-specific skills. Therefore, our analysis in this subsection suggests another reason to expect more general training in economies with low turnover.

The formulation above is also useful in contexts other than merely specific training. For example, s above could be physical capital of the firm. If firms have different levels of physical capital and physical and general human capital are complements, then firms with more physical capital would like to employ workers with more human capital. Suppose that there is one firm that has a higher stock of physical capital than other firms. It would be profitable for this firm to invest in the workers' human capital if physical capital is not perfectly mobile. This conclusion again crucially depends on the existence of some frictions, in this case in the credit market. With perfect markets, a new employer could buy additional physical capital (from the previous employer or another source) and pay the worker his full marginal product, which would prevent the initial employer from recouping training costs. At first sight, this example seems to contradict our general premise that *labor market imperfections* are needed for firm-sponsored investments in general human capital. However, if capital is immobile, then the employer with a larger stock of physi-

---

[11] Related ideas have been discussed in other papers. Stevens (1994) considers skills that are neither completely general nor completely specific and notes that this will mean that workers are unlikely to face a perfect outside labor market for these skills. However, she does not consider the interaction between specific and general skills as a source for firms' investments in general training. Franz and Soskice (1995) discuss the case in which general training is a by-product of specific training; i.e., the complementarity is on the cost side rather than on the output side as in our analysis above. Bishop (1997) points out that individual skills may be general, but the particular mix of these general skills used by any single employer could be firm-specific.

cal capital has monopsony power over the human capital of the worker, which is the source of the distortion in the wage structure. In other words, the imperfection in the capital market spills over into the labor market.

Finally, note that poaching is not a problem here. Because of firm-specific skills, outside firms could not offer higher wages to trained workers and make positive profits.

D. Efficiency Wages

Our analysis in subsection B considered asymmetric information between the current employer of the worker and other firms. Another important asymmetry of information exists between the worker and the firm. Principal-agent, efficiency wage, and personnel economics literatures analyze how the structure of wages can be designed to avoid adverse selection and encourage effort (see, e.g., Weiss 1990; Lazear 1995). Incentive compatibility constraints in these models often distort the structure of wages, which we illustrate here with a simple example.

Suppose that the firm invests in general training in the first period. In the second period, it chooses what wage to offer to the worker, but there is a moral hazard problem that requires the firm to pay an efficiency wage. Either the worker can exert effort at cost e and produce $f(\tau)$, where, as before, τ is general human capital, or he exerts no effort and produces nothing. If e or a variable highly correlated with e were contractible, there would be no moral hazard.[12] Instead, a worker who exerts no effort has a probability p of getting caught. We assume that both the firm and the worker are risk-neutral, and there is a limited liability constraint, so that the worker cannot be paid a negative wage. Finally, to simplify the analysis, we assume that the firm has all the bargaining power and that there are no other reasons for a separation (i.e., $q = 0$).

Since a worker caught shirking will receive zero, when he shirks he saves the effort cost, e, but risks losing his wage with probability p. The incentive compatibility condition to exert effort is therefore $w - e \geq (1 - p)w$. The firm, trying to minimize costs, would choose $w = e/p$ if it can. Notice that the incentive compatibility constraint is independent of skill, which creates the necessary distortion in the

---

[12] It is natural that the effort level of the worker is not always observed. Also, in most firms, rather than the output of an individual worker, only the output of a whole division is observed, and this is not easy to use to provide incentives to individual workers.

wage structure.[13] There is also a participation constraint for the worker to be satisfied. We assume that the worker can obtain his net marginal product $f(\tau) - e$ by quitting but would incur some cost $\Delta > 0$, independent of skills, in the process. Therefore, the participation constraint takes the form $w \geq f(\tau) - \Delta$. The important point to notice is that without moral hazard considerations, the presence of Δ does not induce the firm to invest in general skills (i.e., $\hat{\tau} = 0$).

It is clear that the optimal wage structure, which satisfies the incentive and participation constraints above, is

$$w(\tau) = \max\left\{\frac{e}{p}, f(\tau) - \Delta\right\}. \tag{8}$$

The firm then chooses τ to maximize profits $f(\tau) - w(\tau) - c(\tau)$. It should be clear that this distortion will encourage the firm to invest in general training (as long as e/p is not so high as to shut down production). So in general $\hat{\tau}$ will be positive.[14]

Notice at this point a feature that distinguishes this mechanism (and the minimum-wage and union examples that follow) from the others we have discussed. There is no distortion in the external wage structure ($v'(\tau) = f'(\tau)$), but efficiency wages (or minimum wages or unions) distort the internal wage structure. In contrast, in the other examples, labor market frictions distort the external wage structure and, via this channel, compress the internal wage structure. Nevertheless, in the efficiency wage example, when there are moral hazard problems in other firms as well, the external wage structure will also be distorted, but this does not affect our results. In fact, more generally, (8) will take the form $w(\tau) = \max\{e/p, v(\tau)\}$, which continues to be compressed at low wages irrespective of the form of $v(\tau)$. For example, when all other firms in the economy have exactly the same moral hazard problem, we have $v(\tau) = p_w w(\tau) - \Delta$, where p_w is the probability of employment, which may be less than one as a result of unemployment caused by efficiency wages. It is straightfor-

---

[13] The assumption that the incentive compatibility constraint is independent of future job opportunities, and thus of skills, is not crucial. The result holds as long as the constraint induces a relation between wages and skills less steeply sloped than $f(\tau)$. The model by Loewenstein and Spletzer (1998a) has a similar flavor. In their model, firms can commit ex ante to pay a certain wage in the second period in order to reduce turnover. Whenever this constraint is binding for the firm, it has an incentive to invest in the worker's general skills.

[14] The exact level of training depends on the relative positions of the kink in the wage function (8) and τ^*. In particular, let $\bar{\tau}$ be such that $f(\bar{\tau}) - \Delta = e/p$. Then if $\bar{\tau} \leq \tau^*$ and $c(\bar{\tau}) \leq \Delta$, the firm will operate and choose $\hat{\tau} = \bar{\tau}$. If $\bar{\tau} > \tau^*$ and $f(\tau^*) \geq c(\tau^*) + (e/p)$, then the firm will operate and choose $\hat{\tau} = \tau^*$. Finally, if $\bar{\tau} > \tau^*$ and $f(\tau^*) < c(\tau^*) + (e/p)$ or if $\bar{\tau} \leq \tau^*$ and $c(\bar{\tau}) > \Delta$, then the firm will choose not to operate.

ward to see that firms continue to invest in general training, so whether the rest of the economy is subject to moral hazard is not crucial. Also in this case, the compressed external wage structure induces the internal wage structure to be compressed further and increases firms' incentives to sponsor training.

Finally, there will be no poaching in this case either because Δ is a mobility cost or a premium to specific skills, and the worker loses this amount when he changes jobs. Therefore, an outside firm would need to pay a wage that is at least higher than the current wage by Δ, so poaching is not profitable.

E. Minimum Wages and Other Wage Floors

The next two mechanisms we discuss are labor market institutions that create wage distortions. Perhaps the most common intervention in the labor market is the imposition of wage floors, due to minimum wages and high reservation wages caused by unemployment benefits. Minimum wages are relatively low in the United States and the United Kingdom as compared to the higher levels in many continental European economies.

It is well known that the imposition of a minimum wage can never lead to more training when labor markets are competitive (Rosen 1972). Because workers pay for training through lower wages, a minimum wage may prevent the firm from reducing wages enough during the training period. This is the rationale behind the introduction of "training subminima" in many recent U.S. minimum-wage laws.

Now consider a labor market with frictions, where $v(\tau) = f(\tau) - \Delta$, due to a mobility cost unrelated to skill. It is once again important to emphasize that this distortion does not by itself lead to firm-sponsored training because $v(\tau)$ is not distorted. Also suppose that the firm has all the bargaining power ($\beta = 0$) so that $w(\tau) = v(\tau)$.

Next consider a wage floor w_M due to either minimum wages or unemployment benefits. The structure of wages now becomes

$$w(\tau) = \max\{w_M, f(\tau) - \Delta\}, \tag{9}$$

which is kinked at w_M and thus distorted at low levels of τ. The firm then chooses τ to maximize $f(\tau) - w(\tau) - c(\tau)$. Observe that this wage function is identical to (8) in the case of efficiency wages, with w_M replacing e/p. The condition for a positive training level, a distortion in the structure of wages, is satisfied, so equilibrium training is $\hat{\tau} > 0$ as long as the firm chooses to operate.[15] Note that as in the

---

[15] The exact conditions for training are very similar to those given in n. 14. In particular, w_M replaces e/p, and the no-shutdown conditions become $w_M + c(\bar{\tau}) \leq \Delta$ when $\bar{\tau} < \tau^*$ and $2w_M + c(\tau^*) \leq f(\tau^*)$ when $\tau^* \leq \bar{\tau}$, because the firm has to pay the minimum wage w_M in both periods.

case of efficiency wages, if Δ is interpreted as a mobility cost for the worker, there are no poaching opportunities.

Notice the stark contrast of the predictions in this case to those of the standard human capital model. With competitive markets, a minimum wage just below $f(0)$ is detrimental to the accumulation of general human capital because it prevents the worker from taking a wage cut in the first period to compensate the firm for the costs of training. With frictions, in contrast, such a minimum wage could imply $f'(0) > w'(0) = 0$ and induce the firm to invest in general training.[16]

Given the contrast between our results and those based on Becker's theory of general training in which workers bear the costs, it is instructive to look at the empirical evidence regarding the impact of minimum wages on training. The micro evidence is mixed. Leighton and Mincer (1981) and Neumark and Wascher (1998) find negative effects of minimum wages on training, whereas Grossberg and Sicilian (1999) and Acemoglu and Pischke (1998*a*) find no effects of minimum wages. Only our own study uses within-state variation in the minimum wage from one year to the next, which seems the most convincing way of getting at the effect of the minimum wage. This absence of negative effects of minimum wages on training suggests that as well as preventing some workers from financing their own training, minimum wages may also be inducing firm-sponsored training as implied by our theory.

F. Unions

Another important institutional difference across economies is the role played by unions. In Germany and Scandinavian countries, unions are heavily involved in wage determination, whereas in the United States they have traditionally been less prominent, and their importance has been declining. Since unions tend to compress the wage structure among covered workers (Freeman and Medoff 1984), they are relevant for our theory.

---

[16] There are additional results when we consider the case in which workers can also contribute to training (see Acemoglu and Pischke [1998*a*] for details). For example, if there was no minimum wage previously and workers were able to pay for their own training, then the introduction of a minimum wage leads to firm financing of the training. But the training level never goes up and may go down. On the other hand, an increase in a previously binding minimum wage may lead to more investment. In this case, worker financing through lower wages was already impossible, so that the firm paid for training before the minimum-wage increase. Because a higher minimum wage moves the kink in the wage function (9) to the right, the firm will now choose more training unless it was already providing the first-best amount.

We therefore consider union wage setting as an alternative to individual Nash bargains discussed above. A union with N members can set the entire wage structure $w(\tau)$ at the beginning of the period, and then the firm chooses training. Hence, this model is an analogue to the standard monopoly union (right-to-manage) model, except that because of constant returns, all N union members are employed, but the firm's labor demand decision is replaced with the training decision. The firm maximizes profit per worker $\pi(\tau) = f(\tau) - w(\tau) - c(\tau)$ and chooses to shut down if maximum profits are negative.

We start with the simple case in which the union can choose only one wage for all training levels, $w(\tau) = w$. We shall see that the union cannot improve over this situation. Also suppose that the rest of the economy is not unionized and has a wage structure $v(\tau) < f(\tau)$. The union anticipates the behavior of the firm, which can be summarized by the first-order condition $f'(\tau) - w'(\tau) = c'(\tau)$. Since the wage does not vary with skill (i.e., $w(\tau) = w$), the firm will choose first-best training, τ^*.

The union simply maximizes the wage income of workers and has to make sure to obey $\pi(\tau^*) \geq 0$ so that the firm does not shut down. This implies that the union will set w so as to extract all the rents and force the firm down to zero profits. Therefore, the optimal wage is $w^* = f(\tau^*) - c(\tau^*)$.[17]

The reason for first-best training is that the union is choosing the wage structure before the training decisions. The firm invests τ^* because $f'(\tau) > w'(\tau) = 0$; in other words, workers get a fixed payment and the firm is the full residual claimant. This immediately implies that the union cannot do better by choosing a wage schedule that is different from $w(\tau) = w^*$. It is also important to note that if the firm had the option to set the wage structure itself and *commit* to this ex ante, we would not obtain the same results because the union is not only committing to a wage structure but also choosing one that forces the firm to pay less skilled workers more than their marginal product. The firm would never find this profitable without the union.

Next consider the case in which the rest of the economy is unionized, with other unions choosing the same wage policy. Therefore, a worker who quits the firm will either be unemployed (probability $1 - p_w$) or find a job paying w^* (probability p_w). Therefore, $v(\tau) = p_w w^*$. It is clear that for all values of p_w, this does not change the optimal wage policy of the union, which is again to set w^* and ensure first-best training.

[17] As long as $w^* \geq v(\tau^*)$, which we assume to be the case.

This analysis also suggests a new reason why unions may like to compress the wage structure. Such behavior is usually explained with reference to unions' political preferences (e.g., the median union member chooses the wage structure, and the distribution of marginal products may be skewed to the left) or ideological reasons (see, e.g., Freeman and Medoff 1984). Our analysis points out that when unions take into account the impact of the wage structure on the training decisions of firms, there is another reason for choosing a compressed wage structure. It is also interesting to note that in this case there may be room for poaching since $w^* < f(\tau^*)$. So if the firm anticipates that trained workers will be poached by other firms, it would choose not to invest and shut down. Thus unions might have an additional role in preventing mobility of trained workers, which is sometimes suggested as a role of German works councils.

Notice finally that if there were ex ante heterogeneity among covered workers, the union would no longer choose a single wage. However, it can be shown in this case that the union would still choose to compress the wage structure and induce training.

The predictions of our model once again are different from those of the standard theory, where wage compression would reduce workers' investments in general training. In contrast, we predict that by compressing the wage structure, unions may encourage firms to sponsor training programs. The micro evidence is once again mixed. Studies by Duncan and Stafford (1980) and Mincer (1983) based on the Panel Study on Income Dynamics, Lillard and Tan (1992) based on the Current Population Survey, and Barron, Fuess, and Loewenstein (1987) based on the Employment Opportunity Pilot Project (EOPP) find negative effects of union status on training. Barron, Berger, and Black (1997), on the other hand, report insignificant union effects using the EOPP data and find positive effects for formal training in the Small Business Administration survey. Lynch (1992) also finds positive effects for formal training in the NLSY. For the United Kingdom, Booth (1991) reports more training for union workers, and Green (1993) finds more training for unionized workers in small establishments but not in large establishments.

IV. Conclusion

When the wage structure is distorted away from the competitive benchmark and in favor of less skilled workers, firms may want to invest in the general skills of their employees. For this result to hold, workers do not need to be credit constrained. What matters is the form of labor market frictions and institutions. These results contrast with the standard theory based on Becker's seminal work in

which firms would never invest in general skills. We also found that more frictional and regulated labor markets may encourage more firm-sponsored training.

We view the presence of many firm-sponsored general training programs, such as the German apprenticeship system, and the fact that U.S. employers send their workers to vocational and technical training facilities without reducing their wages as evidence that the forces we emphasize are present. Also, the fact that firms appear to contribute more toward general skills training in Europe and Japan, which have more regulated and frictional markets and more distorted wage structures, is in line with our approach. Future empirical work should test the more micro-level implications that follow from our analysis and contrast them with those of the standard theory.

This paper also has implications for the interpretation of empirical results on the returns to training (e.g., Lynch 1992). Wage returns to training reflect the total increase in productivity only if labor markets are competitive. Our work predicts that, whenever employers pay for training, true returns will exceed wage returns, which are often estimated to be quite large already.

We have discussed a number of reasons why wages may be compressed, but our list is by no means exhaustive. Lazear (1989) argues that pay compression may arise so as to encourage workers to cooperate, or at least to discourage sabotaging their coworkers. Optimal pay compression may also arise when workers can direct their effort between different tasks; the output of only some tasks is easily measured whereas others have an effect on firm profits that is harder to detect. In this case, it may be optimal not to reward observed performance differences, which again compresses the structure of wages (e.g., Holmstrom and Milgrom 1991; Baker 1992). Any of these and other reasons for wage compression may also encourage firms to invest in training. In future work, the link between these stories and training can be more carefully derived, yielding empirical predictions to determine which sources of wage compression, if any, are important in encouraging firm-sponsored training.

Finally, an important development in the theory of contracts in recent years has been the literature on incomplete contracts and property rights (e.g., Grossman and Hart 1986; Hart and Moore 1990). As in the earlier papers by Hashimoto (1979, 1981), Hashimoto and Yu (1980), and Grout (1984), this literature focuses on relationship- or asset-specific investments and analyzes how property rights and other features of organizations can be designed to maximize efficiency. By its nature, this literature has been partial equilibrium. Our analysis was a first attempt at investigating how market structure can turn general skills into effectively relationship-specific

skills, but we have ignored how organizational forms or property rights *within firms* matter for these investments. Combining these two types of analyses may yield new insights into thinking about organizations and markets. For example, how do organizational choices vary with the extent of market frictions? Is it beneficial to make skills more specific? Do different forms of organizations lead to different paths of training, productivity, and wage growth?

Appendix

Dynamic Version of the Search Model

Consider a continuous-time infinite horizon version of the model of Section IIIA. Namely, each worker is matched with a firm, and the firm decides whether and how much to invest in the general skills of the worker. The worker has no funds and cannot commit to a lower wage in the future in return for training now. The productivity of a worker who receives training τ is $f(\tau)$ in every period. For simplicity, training is possible only in period $t = 0$. Both firms and workers are risk-neutral and discount the future at the rate r. All worker-firm matches come to an end at the exogenous rate q. Also a worker, once unemployed, finds a new firm at the rate p_w, which is independent of his training level, and a firm after losing its worker finds a new worker at the rate p_f. We set the unemployment benefit to $b = 0$ to simplify the expressions. The worker that the firm finds will be a random draw from the pool of unemployed workers, irrespective of the value of training. So workers with different levels of training have the same probability of getting a job.

Suppose that all workers have training $\bar{\tau}$, and consider a worker with training τ. Then the value of being employed for this worker as a function of his training level τ, $J^E(\tau)$, is

$$rJ^E(\tau) = w(\tau) + q[J^U(\tau) - J^E(\tau)],$$

where $J^U(\tau)$ is the present discounted value of being unemployed for a worker of training τ. This equation is a standard dynamic programming equation (see, e.g., Pissarides 1990). The worker gets $w(\tau)$ every instant he is with the firm and loses his job at the flow probability q, in which case he gets J^U and loses J^E. In turn we have

$$rJ^U(\tau) = p_w[J^E(\tau) - J^U(\tau)].$$

And for the firm, the value of employing a worker with training τ is

$$rJ^F(\tau) = f(\tau) - w(\tau) + q[J^V - J^F(\tau)],$$

and the value of having an unfilled vacancy is

$$rJ^V = p_f[J^F(\bar{\tau}) - J^V].$$

Nash bargaining in this context implies that the present discounted values should be shared. Therefore, $w(\tau)$ will be chosen so as to maximize

$$[J^E(\tau) - J^U(\tau)]^\beta [J^F(\tau) - J^V]^{1-\beta}.$$

This gives a standard wage rule:

$$w(\tau) = \beta f(\tau) + (1 - \beta) r J^U(\tau) - \beta r J^V$$

or, with substitution for $r J^U(\tau)$,

$$w(\tau) = \frac{(p_w + r + q)[\beta f(\tau) - \beta r J^V]}{r + q + \beta p_w}.$$

Now in period $t = 0$, since the worker is credit constrained and cannot invest in training, the firm will maximize

$$J^F(\tau) - c(\tau) \tag{A1}$$

by choosing training τ and taking the training level of all other workers, $\bar{\tau}$, as given. The term $\bar{\tau}$ influences only J^V, which is in turn independent of the value of τ. So the level of $\bar{\tau}$ does not influence the choice of τ. For this reason, the first-order condition of (A1) takes the simple form

$$(1 - \beta) f'(\hat{\tau}) = (r + q + \beta p_w) c'(\hat{\tau}).$$

Since $c'(0) = 0$, for all $\beta < 1$ and $r + q + \beta p_w < \infty$, the firm will choose $\hat{\tau} > 0$. Since all other firms are solving a similar problem, we also have $\bar{\tau} = \hat{\tau}$ and a unique symmetric equilibrium.

The reason why $\beta < 1$ is necessary for firm-sponsored training is familiar from the text. However, the second condition is interesting. First, it requires that $r < \infty$; thus the future needs to feature in the calculations. A value of $q < \infty$ is also required, which means that the worker should not be leaving the firm for sure. Finally, $p_w < \infty$ is necessary. In fact, $p_w \to \infty$ is the case of perfectly competitive labor markets: the worker finds an employer immediately. Therefore, this last requirement reiterates that labor market imperfections are necessary for firms to invest in the general skills of their workers. Moreover, it is clear that as p_w increases, there is less investment in training. Since steady-state unemployment in this economy is equal to $u = q/(q + p_w)$, this implies that higher unemployment is associated with more investment in training. The reason is that a higher rate of unemployment leads to a more distorted wage structure by reducing the outside option of more skilled workers.

References

Abowd, John, and Lemieux, Thomas. "The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada." *Q.J.E.* 108 (November 1993): 983–1014.

Acemoglu, Daron. "Training and Innovation in an Imperfect Labour Market." *Rev. Econ. Studies* 64 (July 1997): 445–64.

Acemoglu, Daron, and Pischke, Jörn-Steffen. "Minimum Wages and On-the-Job Training." Manuscript. Cambridge: Massachusetts Inst. Tech., Dept. Econ., 1998. (*a*)

———. "Why Do Firms Train? Theory and Evidence." *Q.J.E.* 113 (February 1998): 79–119. (*b*)

Altonji, Joseph G., and Shakotko, Robert A. "Do Wages Rise with Job Se-
    niority?" *Rev. Econ. Studies* 54 (July 1987): 437–59.
Altonji, Joseph G., and Williams, Nicolas. "Do Wages Rise with Job Senior-
    ity? A Reassessment." Working Paper no. 6010. Cambridge, Mass.: NBER,
    April 1997.
Autor, David. "Why Do Temporary Help Firms Provide Free General Skills
    Training?" Manuscript. Cambridge, Mass.: Harvard Univ., Kennedy
    School Government, October 1998.
Baker, George P. "Incentive Contracts and Performance Measurement."
    J.P.E. 100 (June 1992): 598–614.
Barron, John M.; Berger, Mark C.; and Black, Dan A. *On-the-Job Training.*
    Kalamazoo, Mich.: Upjohn Inst. Employment Res., 1997.
Barron, John M.; Fuess, Scott M., Jr.; and Loewenstein, Mark A. "Further
    Analysis of the Effect of Unions on Training." *J.P.E.* 95 (June 1987): 632–
    40.
Becker, Gary S. *Human Capital: A Theoretical and Empirical Analysis, with Spe-
    cial Reference to Education.* New York: Columbia Univ. Press (for NBER),
    1964; 3d ed. Chicago: Univ. Chicago Press (for NBER), 1993.
Bishop, John H. "The Recognition and Reward of Employee Perfor-
    mance." *J. Labor Econ.* 5, no. 4, pt. 2 (October 1987): S36–S56.
———. "The Impact of Previous Training on Productivity and Wages." In
    Training and the Private Sector: International Comparisons, edited by Lisa M.
    Lynch. Series in Comparative Labor Markets. Chicago: Univ. Chicago
    Press (for NBER), 1994.
———. "What We Know about Employer-Provided Training: A Review of
    the Literature." In *Research in Labor Economics,* vol. 16, edited by Solomon
    W. Polachek. Greenwich, Conn.: JAI, 1997.
Blanchflower, David G.; Oswald, Andrew J.; and Sanfey, Peter. "Wages,
    Profits, and Rent-Sharing." *Q.J.E.* 111 (February 1996): 227–51.
Blau, Francine D., and Kahn, Lawrence M. "International Differences in
    Male Wage Inequality: Institutions versus Market Forces." *J.P.E.* 104 (Au-
    gust 1996): 791–837.
Blinder, Alan S., and Krueger, Alan B. "Labor Turnover in the USA and
    Japan: A Tale of Two Countries." *Pacific Econ. Rev.* 1 (June 1996): 27–
    57.
Booth, Alison L. "Job-Related Formal Training: Who Receives It and What
    Is It Worth?" *Oxford Bull. Econ. and Statis.* 53 (August 1991): 281–94.
Chang, Chun, and Wang, Yijiang. "Human Capital Investment under Asym-
    metric Information: The Pigovian Conjecture Revisited." *J. Labor Econ.*
    14 (July 1996): 505–19.
Duncan, Greg J., and Stafford, Frank P. "Do Union Members Receive Com-
    pensating Wage Differentials?" *A.E.R.* 70 (June 1980): 355–71.
Dustmann, Christian, and Meghir, Costas. "Wages, Experience and Senior-
    ity." Manuscript. London: Univ. Coll. London, Dept. Econ., 1997.
Edin, Per-Anders, and Topel, Robert H. "Wage Policy and Restructuring:
    The Swedish Labor Market since 1960." In *The Welfare State in Transition:
    Reforming the Swedish Model,* edited by Richard B. Freeman, Robert H.
    Topel, and Birgitta Swedenborg. Chicago: Univ. Chicago Press (for
    NBER), 1997.
Franz, Wolfgang, and Soskice, David. "The German Apprenticeship Sys-
    tem." In *Institutional Frameworks and Labor Market Performance: Comparative
    Views on the U.S. and German Economies,* edited by Friedrich Buttler et al.
    London: Routledge, 1995.

Freeman, Richard B., and Medoff, James L. *What Do Unions Do?* New York: Basic Books, 1984.

Green, Francis. "The Impact of Trade Union Membership on Training in Britain." *Appl. Econ.* 25 (August 1993): 1033–43.

Grossberg, Adam J., and Sicilian, Paul. "Minimum Wages, On-the-Job Training, and Wage Growth." *Southern Econ. J.* 65 (January 1999).

Grossman, Sanford J., and Hart, Oliver D. "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration." *J.P.E.* 94 (August 1986): 691–719.

Grout, Paul A. "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach." *Econometrica* 52 (March 1984): 449–60.

Harhoff, Dietmar, and Kane, Thomas J. "Is the German Apprenticeship System a Panacea for the U.S. Labor Market?" *J. Population Econ.* 10 (May 1997): 171–96.

Hart, Oliver D., and Moore, John. "Property Rights and the Nature of the Firm." *J.P.E.* 98 (December 1990): 1119–58.

Hashimoto, Masanori. "Bonus Payments, On-the-Job Training, and Lifetime Employment in Japan." *J.P.E.* 87, no. 5, pt. 1 (October 1979): 1086–1104.

———. "Firm-Specific Human Capital as a Shared Investment." *A.E.R.* 71 (June 1981): 475–82.

Hashimoto, Masanori, and Yu, Ben T. "Specific Capital, Employment Contracts, and Wage Rigidity." *Bell J. Econ.* 11 (Autumn 1980): 536–49.

Holmstrom, Bengt, and Milgrom, Paul. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *J. Law, Econ., and Organization* 7 (special issue, 1991): 24–52.

Jovanovic, Boyan. "Job Matching and the Theory of Turnover." *J.P.E.* 87, no. 5, pt. 1 (October 1979): 972–90.

Katz, Eliakim, and Ziderman, Adrian. "Investment in General Training: The Role of Information and Labour Mobility." *Econ. J.* 100 (December 1990): 1147–58.

Krueger, Alan B. "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984–1989." *Q.J.E.* 108 (February 1993): 33–60.

Lazear, Edward P. "Pay Equality and Industrial Politics." *J.P.E.* 97 (June 1989): 561–80.

———. *Personnel Economics.* Cambridge, Mass.: MIT Press, 1995.

Leighton, Linda, and Mincer, Jacob. "The Effects of Minimum Wages on Human Capital Formation." In *The Economics of Legal Minimum Wages*, edited by Simon Rottenberg. Washington: American Enterprise Inst. Public Policy Res., 1981.

Lillard, Lee A., and Tan, Hong W. "Private Sector Training: Who Gets It and What Are Its Effects?" In *Research in Labor Economics*, vol. 13, edited by Ronald G. Ehrenberg. Greenwich, Conn.: JAI, 1992.

Lindbeck, Assar, et al. "Options for Economic and Political Reform in Sweden." *Econ. Policy* 18 (October 1993): 219–46.

Loewenstein, Mark A., and Spletzer, James R. "Dividing the Costs and Returns to General Training." *J. Labor Econ.* 16 (January 1998): 142–71. (*a*)

———. "General and Specific Training: Evidence from the NLSY." Manuscript. Washington: Bur. Labor Statis., 1998. (*b*)

———. "Formal and Informal Training: Evidence from the NLSY." In *Research in Labor Economics*, vol. 18, edited by Solomon W. Polachek. Greenwich, Conn.: JAI, 1999, in press.

Lynch, Lisa M. "Private-Sector Training and the Earnings of Young Workers." *A.E.R.* 82 (March 1992): 299–312.

Mincer, Jacob. "Union Effects: Wages, Turnover, and Job Training." In *New Approaches to Labor Unions*, edited by Joseph D. Reid, Jr. Research in Labor Economics, suppl. 2. Greenwich, Conn.: JAI, 1983.

Neumark, David, and Wascher, William. "Minimum Wages and Training Revisited." Working Paper no. 6651. Cambridge, Mass.: NBER, July 1998.

OECD. *Employment Outlook*. Paris: OECD, 1993, 1994.

Pissarides, Christopher A. *Equilibrium Unemployment Theory*. Cambridge: Blackwell, 1990.

Ritzen, Jozef M. M., and Stern, David. "Introduction and Overview." In *Market Failure in Training? New Economic Analysis and Evidence on Training of Adult Employees*, edited by David Stern and Jozef M. M. Ritzen. Berlin: Springer Verlag, 1991.

Rosen, Sherwin. "Learning and Experience in the Labor Market." *J. Human Resources* 7 (Summer 1972): 326–42.

Ryan, Paul. "The Costs of Job Training for a Transferable Skill." *British J. Indus. Relations* 18 (November 1980): 334–52.

Stevens, Margaret. "A Theoretical Model of On-the-Job Training with Imperfect Competition." *Oxford Econ. Papers* 46 (October 1994): 537–62.

Topel, Robert H. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *J.P.E.* 99 (February 1991): 145–76.

Topel, Robert H., and Ward, Michael P. "Job Mobility and the Careers of Young Men." *Q.J.E.* 107 (May 1992): 439–79.

von Bardeleben, Richard; Beicht, Ursula, and Fehér, Kálmán. *Betriebliche Kosten und Nutzen der Ausbildung: Repräsentative Ergebnisse aus Industrie, Handel und Handwerk*. Berichte zur beruflichen Bildung Heft 187. Bielefeld: Bertelsmann, 1995.

Weiss, Andrew. *Efficiency Wages: Models of Unemployment, Layoffs, and Wage Dispersion*. Princeton, N.J.: Princeton Univ. Press, 1990.