

Appendices

A The payoff to trainees who quit

This appendix sets out a game for the hiring process with hiring at discrete time intervals that is consistent with the employment model in the main text. It shows that the expression in (7) for the remaining lifetime earnings of a trainee who quits is a perfect Bayesian equilibrium payoff of a trainee who quits as the time interval between hiring dates goes to zero.

For simplicity, consider 1 worker and m firms, each with one vacancy, plus the training firm that hires and trains the worker at date 0. Divide the worker's working lifetime $[0, T]$ into a large number N of equally spaced intermediate hiring dates at which job offers can be made, the interval between them being $dt = T/(N + 1)$. The first hiring date at which a non-training firm can make an offer is thus $t = dt$, the training firm being the incumbent employer at this date. The worker works continuously between hiring dates at the wage agreed with the current employer at the most recent hiring date. She quits at the next hiring date for exogenous reasons with a probability that corresponds to the rate ρ during a period of employment. Thus, at each hiring date, the probability the worker quits no matter what offer the current employer makes is $1 - e^{-\rho dt}$. Whatever the reason for a quit, the new employer incurs a once-off hiring cost $\varepsilon > 0$.

At each date ndt at which job offers can be made, the following stage game is played. First, at stage $n.1$, Nature determines whether the worker decides to quit for exogenous reasons (an event that has probability $1 - e^{-\rho dt}$) and informs the current employing firm of the worker's type. At stage $n.2$, the m firms with vacancies each make a simultaneous wage offer to the worker. Each firm knows what wage the worker has received at each previous stage, that is, the worker's wage history. With w_t denoting the wage at t , the wage history at date $t = ndt$ for some n is denoted by the $n - 1$ dimensional vector of real numbers $w^t = (w_0, w_{dt}, \dots, w_{t-dt})$. It does not know the worker's type.²³ At stage $n.3$, the currently employing firm makes a counter-offer to the worker. The firm knows the worker's wage history, the worker's type, and the offers made at stage $n.2$. Finally, at stage $n.4$, the worker either accepts the wage offer of one firm or accepts no offer. If accepting no offer, the worker leaves the labour market permanently and receives a payoff at the rate $u < w^0 + \gamma g(c)$ for subsequent stages. If the worker is quitting for exogenous reasons, the offer accepted must be from an outside firm. Otherwise it can be from any firm. At this stage, the worker knows her own wage history and the offers made to her at stages $n.3$ and $n.4$. The stage game is summarised in Table 3.

For describing payoffs and strategies in the game, we use the following notation. Since all firms know the worker's level of training c , we can simplify notation by denoting the productivity of a better worker (type $\bar{\gamma}$) by $\bar{z} (= w^0 + \bar{\gamma} g(c))$, that of a less good worker (type γ) by $\underline{z} (= w^0 + \gamma g(c))$. The flow payoff to a firm from employing a worker of type $z \in \{\underline{z}, \bar{z}\}$ at wage w_t from t to $t + dt$ is $(z - w_t) dt$, that from not employing a worker is zero. The flow payoff to the worker is $w_t dt$. The structure of the game, and the payoffs conditional on worker type, are common knowledge.

²³This formulation requires that a firm forgets the worker's type when the worker quits its employment. This simplifies the formal model by ensuring that no more than one firm knows the worker's type at any hiring date. We discuss implications of relaxing this requirement later.

	Player to move	Moves	Information Set
Stage n.1	Nature	1. Determines whether worker quits for exogenous reasons 2. Informs employing firm of worker's type	
Stage n.2	Firms with vacancies (simultaneously)	Offer wage to allocated worker	Worker's training level & wage history
Stage n.3	Firm with worker	Counter-offer wage to incumbent worker	1. Worker's training level & wage history 2. Worker's type 3. Offers made to worker at stage n.2
Stage n.4	Worker	Accept no offer (& quit market) or offer of 1 firm from: 1. outside offers, if quitting for exogenous reasons 2. all offers, otherwise	1. Own training level & wage history 2. Offers made at stage n.2 and n.3

Table 3: Stage game at stage n

Strategies in the game can be described as follows. The strategy component at t for each firm $f \in \{1, \dots, m+1\}$ conditional on having a vacancy at t is a mapping from the wage history w^t to a real number that corresponds to a wage offer, denoted $w_t^f(w^t)$. That for each firm $i \in \{1, \dots, m+1\}$ conditional on employing the worker at $t-dt$ is a mapping from the worker's type z , the wage history w^t , and the wage offers by firms with vacancies at t which (with some abuse of notation) we denote $\{w_t^f, f \neq i\}$ to a real number that again corresponds to a wage offer, denoted $\hat{w}_t^i(z, w^t, \{w_t^f, f \neq i\})$. The strategy for the worker at t employed by firm i at $t-dt$ is a mapping from the wage history w^t , the offers by firms with vacancies at t $\{w_t^f, f \neq i\}$, and the counter-offer by firm i , denoted (again with some abuse of notation) \hat{w}_t^i , to the set of integers $\{0, 1, \dots, m+1\}$ if not quitting for exogenous reasons and to the set of integers $\{0, 1, \dots, i-1, i+1, \dots, m+1\}$ if quitting for exogenous reasons. The integer chosen identifies the firm whose offer is accepted, with 0 denoting that no offer is accepted and the worker is leaving the labour market. We denote the worker's strategy component at t when currently employed by firm i by $a_t^i(w^t, \{w_t^f, f \neq i\}, \hat{w}_t^i)$, $i \in \{1, \dots, m+1\}$. The complete strategy space for all players is denoted \mathcal{S} , a profile of strategies (one for each player) by $S \in \mathcal{S}$.

To demonstrate the result in (7), we show that the wage path $w_t = \underline{z}$ for all $t \in (0, T]$ is a perfect Bayesian equilibrium for the continuous version of this game (as $dt \rightarrow 0$) as the number of firms with vacancies $m \rightarrow \infty$ and the hiring cost $\varepsilon \rightarrow 0$. For this demonstration, we use the following notation. Let $w_t^+ = (w_{t+dt}, w_{t+2dt}, \dots, w_{T-dt})$ be a wage path after t . Let $\Pi_t^*(z, w_t, w_t^+)$ denote the expected profits a firm receives from t on from having in post a worker of type $z \in \{\underline{z}, \bar{z}\}$ when the wage path (w_t, w_t^+) is followed from t on and the worker quits only for exogenous reasons. Also let $\bar{\Pi}_t(w_t, w_t^+)$ denote the expected present discounted value of profits for a firm that starts period t with a vacancy in an equilibrium with future wage path (w_t, w_t^+) in which the worker quits only for exogenous reasons and the firm's belief is that the worker is type \bar{z} with probability p . Note from the definition of w_t^+ that $w_t^+ = (w_{t+dt}, w_{t+2dt}^+)$, so we can use $\bar{\Pi}_{t+dt}(w_t^+)$ and $\bar{\Pi}_{t+dt}(w_{t+dt}, w_{t+dt}^+)$ interchangeably. Then

$\bar{\Pi}_t(w_t, w_t^+)$ can be written

$$\bar{\Pi}_t(w_t, w_t^+) = \max \left\{ e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+), \frac{1 - e^{-\rho dt}}{m} [p \Pi_t^*(\bar{z}, w_t, w_t^+) + (1 - p) \Pi_t^*(\underline{z}, w_t, w_t^+) - \varepsilon] \right. \\ \left. + \left(1 - \frac{1 - e^{-\rho dt}}{m} \right) e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+) \right\}, \text{ for } t \in \{dt, 2dt, \dots, T - dt\}, \quad (45)$$

$$= e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+) + \frac{1 - e^{-\rho dt}}{m} \max \left\{ 0, [p \Pi_t^*(\bar{z}, w_t, w_t^+) + (1 - p) \Pi_t^*(\underline{z}, w_t, w_t^+) - \varepsilon] - e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+) \right\}, \text{ for } t \in \{dt, 2dt, \dots, T - dt\}, \quad (46)$$

with $\bar{\Pi}_T(w_{T-dt}^+) = 0$. The first term in braces in (45) is the expected profits from having a vacancy at $t + dt$, namely $\bar{\Pi}_{t+dt}(w_t^+)$, discounted back to t . A firm with a vacancy cannot do worse than this because it can always offer a wage that will never be accepted, earn zero profit in period t and then retain the vacancy until $t + 1$. The second term in braces in (45) is constructed as follows. With probability $1 - e^{-\rho dt}$ the worker will quit at t for exogenous reasons. With m outside firms all bidding the same highest wage w_t , each has probability $1/m$ of acquiring the worker if the worker quits for exogenous reasons. With probability p the worker has high productivity \bar{z} yielding expected future profit $\Pi_t^*(\bar{z}, w_t, w_t^+)$, with probability $1 - p$ low productivity \underline{z} yielding expected future profit $\Pi_t^*(\underline{z}, w_t, w_t^+)$. In either case, the successful hiring firm incurs the hiring cost ε . The remaining term consists of the probability of not successfully hiring the worker multiplied by the expected profits from having a vacancy at $t + dt$ constructed as before. Since the game ends at T , it is necessarily the case that $\bar{\Pi}_T(w_{T-dt}^+) = 0$. Note that, in evaluating $\bar{\Pi}_t(w_t, w_t^+)$ and $\Pi_t^*(z, w_t, w_t^+)$, we do not need to keep track of the number of firms with vacancies that actually bid w_t . Consider $\bar{\Pi}_t(w_t, w_t^+)$. If $\bar{\Pi}_t(w_t, w_t^+) > e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+)$, it is necessarily strictly optimal for all m to do so. If, on the other hand, $\bar{\Pi}_t(w_t, w_t^+) = e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+)$, it makes no difference to the profit how many do so. Now consider $\Pi_t^*(z, w_t, w_t^+)$. We can write $\Pi_t^*(z, w_t, w_t^+)$ as

$$\Pi_t^*(z, w_t, w_t^+) = \sum_{k=1}^{(T-t)/dt} e^{-(r+\rho)(T-t-kdt)} \left\{ (z - w_{T-kdt}) \right. \\ \left. + e^{\rho dt} e^{-rdt} (1 - e^{-\rho dt}) \bar{\Pi}_{T-(k-1)dt}(w_{T-(k-1)dt}, w_{T-(k-1)dt}^+) \right\}, \\ \text{for } t \in \{dt, 2dt, \dots, T - dt\}. \quad (47)$$

This is constructed as follows. With probability $e^{-\rho(T-t-kdt)}$, the worker has not quit by $T - kdt$ so the firm receives profit $z - w_{T-kdt}$ for that period, discounted by the discount factor $e^{-r(T-t-kdt)}$. With probability $e^{-\rho(T-t-(k+1)dt)} (1 - e^{-\rho dt})$, the worker has not quit by $T - (k + 1)dt$ but quits at $T - kdt$. In that event, the firm starts period $T - (k - 1)dt$ without a worker with expected payoff $\bar{\Pi}_{T-(k-1)dt}(w_{T-(k-1)dt}, w_{T-(k-1)dt}^+)$ that is discounted by the factor $e^{-r(T-t-(k-1)dt)}$. Thus, since we do not need to keep track of the number of bidding firms in $\bar{\Pi}_t(w_t, w_t^+)$, we do not need to do so in

$\Pi_t^*(z, w_t, w_t^+)$ either.

Next, define $\underline{w}(w_t^+)$ and $\bar{w}(w_t^+)$ by

$$\underline{w}(w_t^+) = \sup w \text{ such that } \Pi_t^*(\underline{z}, w, w_t^+) \geq e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+), \text{ for } t \in \{dt, 2dt, \dots, T - dt\}, \quad (48)$$

$$\bar{w}(w_t^+) = \sup w \text{ such that } \Pi_t^*(\bar{z}, w, w_t^+) \geq e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+), \text{ for } t \in \{dt, 2dt, \dots, T - dt\}. \quad (49)$$

These are the highest wages a firm would be prepared to pay to retain a worker of type \underline{z} and \bar{z} respectively at t given the future wage path w_t^+ . Since $\bar{z} > \underline{z}$, it follows from (47) that $\bar{w}(w_t^+) > \underline{w}(w_t^+)$. Also, let $w^0(w_t^+)$ denote the lowest wage the worker will accept rather than leave the market given the future wage path w_t^+ . Finally on notation, let $p^f(w^t)$ denote the probability assessment of firm f with a vacancy that a worker with wage history w^t is a better worker, that is, has productivity \bar{z} rather than \underline{z} .

To demonstrate that the specified wage path is a perfect Bayesian equilibrium, we use strategies that consist of the following responses.

Definition 1 *The response of a firm with a vacancy at t is Bertrand if it offers the worker the highest wage at which the expected present value of its future profits is non-negative.*

Definition 2 *Suppose w_t^+ is an equilibrium wage path from $t + dt$ on. An employing firm's response at t is matching if, at stage $n.3$ for $n = t$, the firm offers a worker of type z : (1) a wage w that matches the higher of $w^0(w_t^+)$ and the best outside offer if $\Pi_t^*(z, w, w_t^+) \geq e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+)$; (2) any wage strictly less than w otherwise.*

Definition 3 *Suppose w_t^+ is an equilibrium wage path from $t + dt$ on. The worker's response at t is myopic if, at stage $n.4$ for $n = t$, the worker accepts the highest wage w from among those available if $w \geq w^0(w_t^+)$ and no offer if $w < w^0(w_t^+)$.*

Definition 4 *The worker's response at t is non-discriminating if the following conditions hold. (1) If not quitting for exogenous reasons at t , the worker stays with the current employer if the current employer has matched the best outside offer. (2) When quitting at t , the worker chooses a new employer by randomising with equal probabilities between the best equal outside offers.*

This last response is non-discriminating in the sense that the worker treats equally all firms with vacancies that make equally good offers.

We next show that strategies consisting of the Bertrand, matching and myopic/non-discriminating responses for, respectively, firms with vacancies, the employing firm, and the worker form a perfect Bayesian equilibrium for off-equilibrium-path beliefs that are reasonable in the following sense.

Definition 5 *Suppose the worker has wage path (w_0, \dots, w_{T-dt}) in a perfect Bayesian equilibrium. The beliefs of firm f with a vacancy in period $t + 1$ are reasonable if (1) for a wage at t of $w' < w_t$, $p^f(w^t, w') \leq p^f(w^t, w_t)$, and (2) for a wage at t of $w' > w_t$, $p^f(w^t, w') \geq p^f(w^t, w_t)$.*

For stating the result, it is useful to define the function $G(w, w_t^+)$ by

$$G_t(w, w_t^+) = p\Pi_t^*(\bar{z}, w_t, w_t^+) + (1 - p)\Pi_t^*(\underline{z}, w_t, w_t^+) - \varepsilon - e^{-rdt} \bar{\Pi}_{t+dt}(w_t^+),$$

$$\text{for } t \in \{dt, 2dt, \dots, T - dt\}. \quad (50)$$

This is the gain in expected future profits of a firm with a vacancy that acquires, at date t and wage w , a worker who is type \bar{z} with probability p over the expected future profits it would make if it did not acquire the worker at t .

Proposition 5 *Strategies consisting of the Bertrand, matching and myopic/non-discriminating responses for, respectively, firms with vacancies, the employing firm, and the worker form a perfect Bayesian equilibrium with reasonable beliefs for dt sufficiently small. For the wage path w_t in any such equilibrium, the following properties hold for all $t \in \{dt, 2dt, \dots, T - dt\}$:*

1. $p^f(w^t) = p$ for all firms f with vacancies at t .
2. If there exists a $w \in [w^0(w_t^+), \underline{w}(w_t^+)]$ such that $G_t(w, w_t^+) \geq 0$:

$$w_t = \max w \in [w^0(w_t^+), \underline{w}(w_t^+)] \text{ such that } G_t(w, w_t^+) \geq 0. \quad (51)$$

3. If there exists no $w \in [w^0(w_t^+), \underline{w}(w_t^+)]$ such that $G_t(w, w_t^+) \geq 0$:

$$w_t = w^0(w_t^+). \quad (52)$$

Proof. Consider the last hiring period $n = T - dt$. With no future to be concerned about, the unique best response of the worker at stage $n.4$ is clearly the myopic response of accepting the offer with the highest wage w for period $T - dt$ from among those available, or no offer if $w < u$. It cannot do better than by being non-discriminatory. Thus the myopic/non-discriminatory response is a best response at $T - dt$. Also with no future to be concerned about, $\bar{\Pi}_T(w_{T-dt}^+) = 0$. Thus, a best response at stage $n.3$ for the incumbent employer, given that it knows the worker's productivity z , is the matching response of offering w that matches the higher of the worker's reservation wage u and the best outside offer if $\Pi_{T-dt}^*(z, w, w_{T-dt}^+) \geq 0$, or any wage offer less than that w if $\Pi_{T-dt}^*(z, w, w_{T-dt}^+) < 0$.

Consider now the best responses of firms with vacancies making offers to the worker at stage $n.2$ for $n = T - dt$. The only offers consistent with the claimed equilibrium are offers $w \leq \underline{w}(w_{T-dt}^+)$ because otherwise, given the best responses of the incumbent employer and the worker, the worker will certainly receive a wage $w_{T-dt} > \underline{w}(w_{T-dt}^+)$. Suppose firm f with a vacancy, and belief $p^f(w^{T-dt})$ that the worker is type \underline{z} , deviates by offering $w > \underline{w}(w_{T-dt}^+)$. Whatever its beliefs $p^f(w^{T-dt})$ about the worker's type, $w \geq \bar{w}(w_{T-dt}^+)$ is clearly not a best response because $\Pi_{T-dt}^*(z, w, w_{T-dt}^+) \leq 0$ for both \bar{z} and \underline{z} for any $w \geq \bar{w}(w_{T-dt}^+)$ so, given the hiring cost ε , it would certainly achieve a higher payoff by bidding $w < u$ and being sure not to hire the worker. For an offer w such that $\underline{w}(w_{T-dt}^+) < w < \bar{w}(w_{T-dt}^+)$, it follows from the definitions of $\underline{w}(w_t^+)$ and $\bar{w}(w_t^+)$ that the incumbent employer matches w for a \bar{z} type worker but not for a \underline{z} type worker. Thus firm f acquires a type \bar{z} worker if quitting for exogenous reasons (that is, with probability $(1 - e^{-\rho dt}) p^f(w^{T-dt})$) and a type \underline{z} worker whether or not quitting for exogenous reasons (that is, with probability $1 - p^f(w^{T-dt})$). The probability it acquires a worker of either sort (and so incurs the hiring cost ε) is thus $(1 - e^{-\rho dt}) p^f(w^{T-dt}) + (1 - p^f(w^{T-dt})) = 1 - e^{-\rho dt} p^f(w^{T-dt})$. With the remaining probability it does not acquire the worker and receives zero flow payoff from $T - dt$ to T , so its payoff from

the deviation is $\hat{\Pi}_{T-dt}^f(w^{T-dt}, w, w_{T-dt}^+)$ given by

$$\begin{aligned} \hat{\Pi}_{T-dt}^f(w^{T-dt}, w, w_{T-dt}^+) &= [(1 - e^{-\rho dt}) p^f(w^{T-dt}) \Pi_{T-dt}^*(\bar{z}, w, w_{T-dt}^+) \\ &\quad + (1 - p^f(w^{T-dt})) \Pi_{T-dt}^*(\underline{z}, w, w_{T-dt}^+) - (1 - e^{-\rho dt}) p^f(w^{T-dt}) \varepsilon]. \end{aligned} \quad (53)$$

Now, it follows from (48) and $\bar{\Pi}_T(w_{T-dt}^+) = 0$ that $\Pi_{T-dt}^*(\underline{z}, w, w_{T-dt}^+) < 0$ for any $w > \underline{w}(w_{T-dt}^+)$. Moreover, as $dt \rightarrow 0$, $(1 - e^{-\rho dt}) \rightarrow 0$, whereas $(1 - e^{-\rho dt}) p^f(w^{T-dt}) \rightarrow (1 - p^f(w^{T-dt})) > 0$ for $p^f(w^{T-dt}) < 1$. It thus follows that

$$\begin{aligned} \lim_{dt \rightarrow 0} \hat{\Pi}_{T-dt}^f(w^{T-dt}, w, w_{T-dt}^+) &= (1 - p^f(w^{T-dt})) \lim_{dt \rightarrow 0} [\Pi_{T-dt}^*(\underline{z}, w, w_{T-dt}^+) - \varepsilon] \\ &< 0, \end{aligned} \quad (54)$$

the strict inequality holding even if $\Pi_{T-dt}^*(\underline{z}, w, w_{T-dt}^+) \rightarrow 0$ as $w \rightarrow \underline{w}(w_{T-dt}^+)$. Thus, there exists $\delta > 0$ such that, for any $dt < \delta$, $\hat{\Pi}_{T-dt}^f(w^{T-dt}, w, w_{T-dt}^+) < 0$ for any deviation $w > \underline{w}(w_{T-dt}^+)$. Any such deviation is not, therefore, a best response for dt sufficiently small.

For the claimed equilibrium $p^f(w^{T-dt}) = p$ for all f . In that case, playing the Bertrand response is a best response by the standard argument for Bertrand games. That involves offering the highest w such that $G_{T-dt}(w, w_{T-dt}^+) \geq 0$ if that results in $u [= w^0(w_{T-dt}^+)] \leq w \leq \underline{w}(w_{T-dt}^+)$ and offering some $w < u$ otherwise. Given the matching and myopic responses of the incumbent employer and the worker, the wage w_{T-dt} actually paid then satisfies Properties 2 and 3 of the proposition. Note that, if any play before $T - dt$ were to result in $p^f(w^{T-dt}) \neq p$, the wage offer by firm f would be non-decreasing in $p^f(w^{T-dt})$.

Now consider $n = T - 2dt$. Given the equilibrium plays at $T - dt$, the only reason the worker might deviate from the myopic response by turning down the highest wage offer for period $T - 2dt$ is if it would alter the beliefs of outside firms at $T - dt$ in such a way as to increase the wage w_{T-dt} . The possible deviations are for the worker to accept a lower offer or no offer ($w_{T-2dt} = 0$). But, given that the wage offer at $T - dt$ by outside firm f is non-decreasing in $p^f(w^{T-dt})$, for reasonable beliefs as defined such a deviation is worse for the present and no better for the future, so it is not a best response. The worker still cannot do better than by being non-discriminatory. Thus the myopic/non-discriminatory response is a best response at $T - 2dt$. Similarly, the current employer's best response continues to be the matching response. If the higher of the worker's reservation wage and the best outside offer is w , if $\Pi_{T-2dt}^*(z, w, w_{T-2dt}^+) \geq e^{-r dt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+)$, and if the current employer offers less than w , it loses the worker and so reduces its payoff, thus doing worse than by matching. If, on the other hand, it offers more than w , that reduces profits in the current period. Moreover, given reasonable beliefs, it does not reduce the wages outside firms will offer for that worker at $T - dt$. Thus the incumbent employer offers no more than w . If $\Pi_{T-2dt}^*(z, w, w_{T-2dt}^+) < e^{-r dt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+)$, only an offer less than w is a best response. But any such offer is a matching response as defined.

Consider now the best responses of firms with vacancies making offers to the worker at stage $n.2$ for $n = T - 2dt$. The only offers consistent with the claimed equilibrium are offers $w \leq \underline{w}(w_{T-2dt}^+)$ because otherwise, given the best responses of the incumbent employer and the worker, the worker will certainly receive a wage $w_{T-2dt} > \underline{w}(w_{T-2dt}^+)$. Suppose firm f with a vacancy, and belief $p^f(w^{T-2dt})$

that the worker is type \underline{z} , deviates by offering $w > \underline{w}(w_{T-2dt}^+)$. Whatever its beliefs $p^f(w^{T-2dt})$ about the worker's type, $w \geq \bar{w}(w_{T-2dt}^+)$ is clearly not a best response because $\Pi_{T-2dt}^*(z, w, w_{T-2dt}^+) \leq e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+)$ for both \bar{z} and \underline{z} for any $w \geq \bar{w}(w_{T-2dt}^+)$ so, given the hiring cost ε , it would certainly achieve a higher payoff by bidding $w < w^0(w_{T-2dt}^+)$ and being sure not to hire the worker. For an offer w such that $\underline{w}(w_{T-2dt}^+) < w < \bar{w}(w_{T-2dt}^+)$, it follows from the definitions of $\underline{w}(w_i^+)$ and $\bar{w}(w_i^+)$ that the incumbent employer matches w for a \bar{z} type worker but not for a \underline{z} type worker. Thus firm f acquires a type \bar{z} worker if quitting for exogenous reasons (that is, with probability $(1 - e^{-\rho dt}) p^f(w^{T-2dt})$) and a type \underline{z} worker whether or not quitting for exogenous reasons (that is, with probability $1 - p^f(w^{T-2dt})$). The probability it acquires a worker of either sort (and so incurs the hiring cost ε) is thus $(1 - e^{-\rho dt}) p^f(w^{T-2dt}) + (1 - p^f(w^{T-2dt})) = 1 - e^{-\rho dt} p^f(w^{T-2dt})$. With the remaining probability it does not acquire the worker and receives zero payoff, so its payoff from the deviation is $\hat{\Pi}_{T-2dt}^f(w^{T-2dt}, w, w_{T-2dt}^+)$ given by

$$\begin{aligned} \hat{\Pi}_{T-2dt}^f(w^{T-2dt}, w, w_{T-2dt}^+) &= [(1 - e^{-\rho dt}) p^f(w^{T-2dt}) \Pi_{T-2dt}^*(\bar{z}, w, w_{T-2dt}^+) \\ &\quad + (1 - p^f(w^{T-2dt})) \Pi_{T-2dt}^*(\underline{z}, w, w_{T-2dt}^+) \\ &\quad - (1 - e^{-\rho dt} p^f(w^{T-2dt})) \varepsilon] + e^{-\rho dt} p^f(w^{T-2dt}) e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+). \end{aligned} \quad (55)$$

Now, it follows from (48) that $\Pi_{T-2dt}^*(\underline{z}, w, w_{T-2dt}^+) < e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+)$ for any $w > \underline{w}(w_{T-2dt}^+)$. Again, as $dt \rightarrow 0$, $(1 - e^{-\rho dt}) \rightarrow 0$, whereas $(1 - e^{-\rho dt} p^f(w^{T-2dt})) \rightarrow (1 - p^f(w^{T-2dt})) > 0$ for $p^f(w^{T-2dt}) < 1$. It thus follows that

$$\begin{aligned} \lim_{dt \rightarrow 0} \hat{\Pi}_{T-2dt}^f(w^{T-2dt}, w, w_{T-2dt}^+) &= \lim_{dt \rightarrow 0} \{e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+) \\ &\quad + (1 - p^f(w^{T-2dt})) [\Pi_{T-2dt}^*(\underline{z}, w, w_{T-2dt}^+) - \varepsilon \\ &\quad - e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+)]\} \end{aligned} \quad (56)$$

$$< \lim_{dt \rightarrow 0} e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+), \quad (57)$$

the strict inequality holding even if $\Pi_{T-2dt}^*(\underline{z}, w, w_{T-2dt}^+) \rightarrow e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+)$ as $w \rightarrow \underline{w}(w_{T-2dt}^+)$. Thus, there exists $\delta > 0$ such that, for any $dt < \delta$, $\hat{\Pi}_{T-2dt}^f(w^{T-2dt}, w, w_{T-2dt}^+) < e^{-rdt} \bar{\Pi}_{T-dt}(w_{T-2dt}^+)$ for any deviation $w > \underline{w}(w_{T-2dt}^+)$. Thus any such deviation gives a lower payoff than making an offer that the worker will never accept, and is therefore not a best response, for dt sufficiently small.

For the claimed equilibrium $p^f(w^{T-2dt}) = p$ for all f . In that case, playing the Bertrand response is a best response by the standard argument for Bertrand games. That involves offering the highest w such that $G_{T-2dt}(w, w_{T-2dt}^+) \geq 0$ if that results in $w^0(w_{T-2dt}^+) \leq w \leq \underline{w}(w_{T-2dt}^+)$ and offering some $w < w^0(w_{T-2dt}^+)$ otherwise. Given the matching and myopic responses of the incumbent employer and the worker, the wage w_{T-2dt} actually paid then satisfies Properties 2 and 3 of the proposition. Note that, if any play before $T - 2dt$ were to result in $p^f(w^{T-2dt}) \neq p$, the wage offer by firm f would be non-decreasing in $p^f(w^{T-2dt})$.

Backwards iteration of the above argument for earlier dates establishes that the myopic and non-discriminating response by the worker, the matching response by the current employer, and the Bertrand response for firms with vacancies are all best responses for each date $t \in \{dt, 2dt, \dots, T - dt\}$, condi-

tional on $p^f(w^t) = p$, and result in a wage path that satisfies Properties 2 and 3 of the proposition.

It is common knowledge at $t = 0$ that the worker is type \bar{z} with probability $p \in (0, 1)$. It follows that, in the first hiring date $t = dt$, $p^f(w^{dt}) = p \in (0, 1)$ for all firms f with a vacancy. Along the equilibrium path, the wage set at this date is $w_{dt} \leq \underline{w}(w_{dt}^+)$ by Properties 2 and 3 of the proposition and the worker quits only for exogenous reasons. Since the probability of exogenous quits is the same for both types, whether or not the worker quits provides no information about type. Thus, Bayesian updating implies $p^f(w^{2dt}) = p \in (0, 1)$ for all firms f with a vacancy. A similar argument applies at each hiring date t . Thus, along an equilibrium path $p^f(w^t) = p$ for all firms f with a vacancy and Property 1 of the proposition holds for all $t \in \{dt, 2dt, \dots, T - dt\}$. ■

This result has an important consequence. It is not just that strategies consisting of the Bertrand, matching and myopic/non-discriminating responses for, respectively, firms with vacancies, the incumbent employer, and the worker form a perfect Bayesian equilibrium. It is also that, in any such equilibrium, a firm with a vacancy at t never bids more for the worker than $\underline{w}(w_t^+)$, the wage that would be matched by the current employer for a worker of the less good type, if the time interval between offers is sufficiently short. The intuition is as follows. When the employing firm plays the matching response and the worker the myopic response, a bidding firm acquires the worker only at a wage for which the worker's future profitability is negative—unless, that is, the worker is quitting for exogenous reasons. For an offer above $\underline{w}(w_t^+)$, it acquires the worker only if of type \underline{z} or if quitting for exogenous reasons. In the former case, the acquiring firm makes a loss in period t and no additional profit subsequently. But as the time interval dt goes to zero, the probability that a worker quits for exogenous reasons becomes negligible. Thus, as long as there is a hiring cost, the acquiring firm necessarily loses by offering more than $\underline{w}(w_t^+)$.

Property 2 in Proposition 5 is the standard condition for an interior solution given the restriction to wage offers at t no more than $\underline{w}(w_t^+)$. Bertrand competition between firms with vacancies at t pushes the wage offers up to the highest level at which a firm that acquires the worker makes at least as much expected profit as from the alternative of making an offer that is refused for sure and then starting the next period dt later with a vacancy. Since, however, it is not an equilibrium to offer $w > \underline{w}(w_t^+)$ at t , firms with vacancies may make strictly positive expected profits from acquiring the worker at the equilibrium wage $\underline{w}(w_t^+)$ because the wage cannot be pushed higher to remove those profits. Formally, there is a discontinuity in the profit function at $\underline{w}(w_t^+)$ because, for any higher wage, the worker always accepts if of type \underline{z} , not just if quitting for exogenous reasons. Property 3 in Proposition 5 recognises that, because of the hiring cost ε , it may not be profitable for firms with vacancies to bid even the reservation wage $w^0(w_t^+)$ for the worker. Then the incumbent employer simply offers $w_t = w^0(w_t^+)$. (That $\underline{z} > u$ rules out the possibility that there is no wage at which employment can continue in the currently employing firm.)

It can in fact be shown (though at considerably greater length) that any wage path in a perfect Bayesian equilibrium with reasonable beliefs and a non-discriminating response by the worker satisfies Properties 1–3 of Proposition 5 for dt sufficiently small. Even if the worker's response is not non-discriminating, any perfect Bayesian equilibrium with reasonable beliefs can, for dt sufficiently small, be described by Properties 1–3 of Proposition 5 with the function $G(\cdot)$ appropriately modified.²⁴

²⁴In the game to which Proposition 5 applies, a firm quit by the worker at a previous hiring date forgets the worker's type.

The expression in (7) in the main text for the remaining lifetime earnings of a trainee who quits is for the limiting case as the interval dt between wage offers and the hiring cost ε go to zero, while the number of firms with vacancies m goes to infinity. In taking these limits, we specify that the limit first be taken as $dt \rightarrow 0$ in order to ensure that dt is always sufficiently small for the results of Proposition 5 to apply.

Definition 6 *A limiting equilibrium wage path is the limit as $dt \rightarrow 0$, $m \rightarrow \infty$ and $\varepsilon \rightarrow 0$ (with limits taken in that order) of a wage path that satisfies Properties 1–3 of Proposition 5.*

Proposition 6 *There is a unique limiting equilibrium wage path. It satisfies $w_t = \underline{w}(w_t^+) = \underline{z}$ for all $t \in (0, T]$.*

Proof. It follows from (46) and (47) that

$$\lim_{m \rightarrow \infty} \bar{\Pi}_t(w_t, w_t^+) = 0, \text{ for all } t \in (0, T], \quad (58)$$

$$\lim_{m \rightarrow \infty} \lim_{dt \rightarrow 0} \Pi_t^*(z, w_t, w_t^+) = \int_t^T e^{-(r+\rho)(\theta-t)} (z - w_\theta) d\theta, \text{ for all } t \in (0, T]. \quad (59)$$

Thus, from (50),

$$\lim_{m \rightarrow \infty} \lim_{dt \rightarrow 0} G_t(w_t, w_t^+) = \int_t^T e^{-(r+\rho)(\theta-t)} [\underline{z} + p(\bar{z} - \underline{z}) - w_\theta] d\theta - \varepsilon, \text{ for all } t \in (0, T]. \quad (60)$$

As $\varepsilon \rightarrow 0$, therefore, the highest wage at t for which $\lim_{m \rightarrow \infty} \lim_{dt \rightarrow 0} G_t(w_t, w_t^+) \geq 0$ for all t approaches \hat{w} defined by

$$\hat{w} = \underline{z} + p(\bar{z} - \underline{z}). \quad (61)$$

By Property 2 of Proposition 5, this is an equilibrium wage if $\hat{w} \leq \underline{w}(w_t^+)$. From (48) and (58), we have that $\underline{w}(w_t^+)$ is set such that

$$\lim_{m \rightarrow \infty} \lim_{dt \rightarrow 0} \Pi_t^*(\underline{z}, \underline{w}(w_t^+), w_t^+) = 0, \text{ for all } t \in (0, T]. \quad (62)$$

Because this must hold for all t , the derivative of the left hand side with respect to t must also be zero for all t . That and (59) imply

$$\underline{z} - \underline{w}(w_t^+) + (r + \rho) \int_t^T e^{-(r+\rho)(\theta-t)} (\underline{z} - w_\theta) d\theta = 0, \text{ for all } t \in (0, T]. \quad (63)$$

What happens if that is not the case? If two firms both know the worker's type at $t + dt$, Bertrand competition results in the worker receiving a wage equal to her true marginal product from $t + dt$ on. If the incumbent employer adopts the matching response at t , it is always worthwhile for the worker to quit at t because a second firm will then learn the worker's type, which results in a higher wage from $t + dt$ on, without the worker incurring a lower wage from t to $t + dt$. With a positive hiring cost, however, firms with vacancies offer a wage at t strictly less than $\underline{w}(w_t^+)$ for at least part of the working lifetime, and for all of it if that lifetime is sufficiently short. The best response of the incumbent employer may then be to more than match the best outside offer so that the worker incurs a short-term loss from quitting. As long as a wage at $t - dt$ no higher than $\underline{w}(w_{t-dt}^+)$ is enough to stop the worker quitting, a wage path satisfying Properties 1–3 of Proposition 5 remains a perfect Bayesian equilibrium. That will not be the case, however, as both the interval between hiring dates and the hiring cost go to zero because then both the difference between $\underline{w}(w_{t-dt}^+)$ and the best outside offer, and the period for which this difference can be paid, also go to zero. We do not explore this possibility further here because it requires moving away from the limiting case as dt and ε go to 0 that is a very convenient simplification for deriving the results in the main text.

It follows that $\underline{w}(w_T^+) = \underline{z}$. But $\bar{z} > \underline{z}$ and we know from Proposition 5 that the wage cannot exceed $\underline{w}(w_t^+)$ for any t . Hence, it must be that $w_T = \underline{z}$. It follows directly from repeating this argument for earlier t that $w_t = \underline{w}(w_t^+) = \underline{z}$ for all t . ■

The intuition behind this result is as follows. As the number of firms with vacancies becomes large and the hiring cost becomes small, firms with vacancies would be prepared to pay up to the average productivity of the two types, $\underline{z} + p(\bar{z} - \underline{z})$, to hire the worker if quitting for exogenous reasons. But the incumbent employer is willing to pay only up to \underline{z} to retain a worker of the less productive type. So, if firms with vacancies offer more than \underline{z} , a less productive worker will always accept, not just if quitting for exogenous reasons, and whichever outside firm acquires the worker makes a loss on that type. That swamps the higher productivity of the more productive type quitting for exogenous reasons. Thus, outside offers never rise above \underline{z} even though any outside firm that acquires the worker at that wage makes a strictly positive profit from doing so.

The expression in (7) for the remaining lifetime earnings of a trainee who quits follows directly from Proposition 6 in view of the definition of $\underline{z} = w^0 + \gamma g(c)$.

B The incentive constraints for $t > 0$

This appendix shows that, given any $W(0, \tau)$ that satisfies (23) for $t = 0$, it is always possible to find a wage path $w(t)$ that satisfies (23) for all $0 < t \leq \tau$ with $w(0) > w^0$ and with the training firm wishing to retain less good trainees up to τ .

First note that, since $w(t)$ is a flow, its value at the isolated point $t = 0$ is negligible relative to its present discounted value over the interval τ . Thus, we can set $w(0) = w^0 + \varepsilon$ for $\varepsilon > 0$ without affecting subsequent values of $w(t)$ for given $W(0, \tau)$. Let the wage over the remainder of the contract period be constant at

$$w = \frac{W(0, \tau)}{\delta(\tau)} \geq [w^0 + \gamma g(c)] - (1 - \gamma) g(c) e^{-\rho\tau} \frac{\beta(T) - \beta(\tau)}{\delta(\tau)}, \quad (64)$$

where the inequality follows from (23) being satisfied at $t = 0$. Then, from (8),

$$W(t, \tau) = \delta(\tau - t) w \quad (65)$$

$$\geq \delta(\tau - t) [w^0 + \gamma g(c)] - (1 - \gamma) g(c) e^{-\rho\tau} \frac{\delta(\tau - t)}{\delta(\tau)} [\beta(T) - \beta(\tau)]. \quad (66)$$

Since, by construction, this satisfies (23) for $t = 0$, it is clear by inspection that it will also satisfy (23) for $0 < t \leq \tau$ if

$$e^{-\rho\tau} \frac{\delta(\tau - t)}{\delta(\tau)} [\beta(T) - \beta(\tau)] \leq e^{-\rho(\tau - t)} [\beta(T - t) - \beta(\tau - t)], \text{ for } 0 < t \leq \tau, \quad (67)$$

or

$$\frac{\delta(\tau - t)}{\delta(\tau)} \leq e^{\rho t} \frac{\beta(T - t) - \beta(\tau - t)}{\beta(T) - \beta(\tau)}, \text{ for } 0 < t \leq \tau, \quad (68)$$

or, using the definitions of $\beta(t)$ and $\delta(t)$ in (1) and (2),

$$\frac{1 - e^{-(r+\rho)(\tau-t)}}{1 - e^{-(r+\rho)\tau}} \leq e^{(r+\rho)t}, \text{ for } 0 < t \leq \tau,$$

or

$$e^{(r+\rho)t} \geq 1, \text{ for } 0 < t \leq \tau.$$

This is certainly true given r and $\rho \geq 0$.

We next show that, at the wage in (64), the training firm wishes to retain even less good trainees up to τ . To see this, note that the training firm would wish to continue the contract after $0 < t \leq \tau$ even for type γ trainees provided

$$\delta(\tau - t) [w^0 + \gamma g(c)] - W(t, \tau) + e^{(r+\rho)t} S \geq 0 \text{ for } 0 < t \leq \tau. \quad (69)$$

Now consider (69) for the total wage bill given by (24) so that (23) holds with equality at $t = 0$ and $W(t, \tau)$ is given by equality in (66). Then (69) can be written

$$e^{(r+\rho)t} S \geq -(1 - \gamma) g(c) e^{-\rho\tau} \frac{\delta(\tau - t)}{\delta(\tau)} [\beta(T) - \beta(\tau)], \text{ for } 0 < t \leq \tau, \quad (70)$$

which, since $\gamma < 1$ and $\beta(T) \geq \beta(\tau)$, is clearly true for any $S \geq 0$. Finally, suppose that (23) holds with strict inequality, in which case any optimal apprenticeship has $W(0, \tau) = \bar{W}$. Then the addition over the present discounted value of wages making (23) hold with equality that is required to make $W(0, \tau) = \bar{W}$ can all be paid to the worker on agreeing to the contract at $t = 0$ and wages for $0 < t \leq \tau$ paid as above. Then clearly (70) is unaffected and continues to hold.

C Training when it is efficient to train all workers

This appendix considers training outcomes when it is efficient to train all workers, that is $k_1^* \geq k_2^*$. Then \bar{W} , the lifetime value of wages that equates demand and supply for workers at initial hiring, can be greater than βw^0 because the supply of workers may be less than the demand for workers by firms wanting to train at the lifetime value of wages βw^0 . A market equilibrium can be defined formally as follows. Let $\bar{k}(\bar{W})$ denote the highest fixed cost of training at which it is profitable to train with a profit maximizing training arrangement when the expected lifetime utility of workers hired in the market at age 0 is \bar{W} . Then \bar{W} is a market equilibrium if it satisfies

$$\bar{k}(\bar{W}) \leq k_2^* \quad (71)$$

$$\bar{W} \geq \beta w^0 \quad (72)$$

$$\text{either } \bar{W} = \beta w^0 \quad \text{or } \bar{k}(\bar{W}) = k_2^*. \quad (73)$$

(k_2^* was defined above as the highest capital cost that needs to be incurred for all workers to be trained.)

C.1 Training without contracts

Suppose \bar{W} is sufficiently high that (10) is not a binding constraint at $t = 0$ for any level of c chosen by a training firm. What then limits how low the wage can go is that the expected utility of joining a training firm given by (9) with $t = 0$ is at least as great as \bar{W} . That is, expected wages from the training firm W must satisfy

$$W \geq \bar{W} - (\beta - \delta) [w^0 + \gamma g(c)]. \quad (74)$$

Substitution of this into the expression for a training firm's expected profits (11) allows those expected profits to be written

$$\Pi^1(k, c, \bar{W}) = [(1 - \gamma)\delta + \gamma\beta]g(c) + \beta w^0 - \bar{W} - (k + c) \quad (75)$$

for which the first-order condition for the optimal level of training c^1 is

$$[\delta + \gamma(\beta - \delta)]g'(c^1) = 1. \quad (76)$$

It follows from (15) and strict concavity of $g(\cdot)$ that $c^1 > c^0$, though still $c^1 < c^*$ defined in (4). The intuition is that, although a firm's profits from training are lower when wages are higher, the worker captures less of the *marginal* returns to additional training. Indeed, since $\beta > \delta$, a higher value of γ actually *increases* the amount of training, which approaches the efficient level c^* as $\gamma \rightarrow 1$. The reason is that more training enables those workers who quit for exogenous reasons to obtain a higher wage in another firm and, because (10) is not binding, the training firm is able to lower its wage as a result. As γ approaches 1, those quitting for exogenous reasons capture all the return on training after they quit, so their new employers receive no external benefit from the training.

If (10) at $t = 0$ is not binding at c^0 but becomes binding at some $c < c^1$, then c is the profit maximizing level of training for training firms to provide. Moreover, for (10) not to be a binding constraint for some training c , it must be that the expected lifetime value of wages firms must pay to hire trainees, \bar{W} , exceeds what trained workers can obtain by quitting, given by (7) for $t = 0$. That is, $\bar{W} \geq \beta [w^0 + \gamma g(c)]$. For $\bar{W} > \beta w^0$ to be an equilibrium, it must be that all workers are trained at that wage. (This is expressed formally by (73).) Since k_2^* is defined as the fixed cost that must be incurred by the marginal training firm when all workers are trained, it follows from (75) that all workers will actually be trained at this wage only if

$$k_2^* \leq [(1 - \gamma)\delta + \gamma\beta]g(c) + \beta w^0 - \bar{W} - c, \quad (77)$$

or, with \bar{W} at the lowest level $\beta [w^0 + \gamma g(c)]$,

$$k_2^* \leq (1 - \gamma)\delta g(c) - c. \quad (78)$$

Thus, if (78) is satisfied for $c = c^0$, the market equilibrium level of training is the higher level that satisfies (78) with equality provided that level is less than c^1 , and c^1 otherwise. The analysis in the main text applies whenever (78) is not satisfied at $c = c^0$. In view of the definition of k_1^* in (5), the condition

for that can be written

$$k_1^* < k_2^* + [\beta g(c^*) - c^*] - (1 - \gamma) \delta g(c^0) + c^0. \quad (79)$$

For higher values of k_1^* , all workers are trained and the amount of training received by each trainee is the lower of c^1 and the amount that gives equality in (78).

C.2 Training contracts

Suppose \bar{W} is sufficiently high that (23) is not a binding constraint at $t = 0$ for any level of c chosen by a training firm. That does not alter the analysis of the optimal training for any given contract, so (18) still applies. What, however, now limits how low the wage can go is that the expected utility of joining a training firm given by (22) with $t = 0$ is at least as great as \bar{W} . That is, expected wages W from the contract must satisfy

$$W \geq \bar{W} - [w^0 + \gamma g(c)] [\beta - \delta(\tau)] - (1 - \gamma) g(c) e^{-\rho\tau} [\beta - \beta(\tau)]. \quad (80)$$

Substitution of this into the expression for a training firm's expected profits (17), when there is no subsidy ($S = 0$) and training length $\tau(c)$ is chosen to induce training c , allows those expected profits to be written

$$\tilde{\Pi}^1(k, c, \bar{W}) = \left[(1 - \gamma) \tilde{\delta}(\tau(c)) + \gamma \beta \right] g(c) + \beta w^0 - \bar{W} - (k + c), \quad (81)$$

where $\tilde{\delta}(\tau)$ is defined in (26). Since, from (28), $\tilde{\delta}(\tau) > \delta$ for $\tau < T$, expected profits from offering a training contract given by (81) are greater than those from not offering a training contract given by (75) for $c < \bar{c}$ defined by (19), and are the same for $c = \bar{c}$, exactly as in the case of $k_1^* < k_2^*$ discussed in the main text. The difference in the case $k_1^* \geq k_2^*$ is that the level of training that is optimal with no contract if (78) is satisfied with strict inequality for $c = \bar{c}$ is not feasible with a training contract, for which the highest feasible level of training is \bar{c} that is attained when $\tau = T$. Moreover, that optimal level yields higher expected profit than $c = \bar{c}$ when no contract is used and, since the expected profit with $c = \bar{c}$ is the same with a contract as without, it also yields higher profit than $c = \bar{c}$ with a contract. Thus the argument in the main text does not guarantee that firms make greater expected profit with a contract than without in this case.

The argument in the main text does, however, apply as long as the level of training chosen without a contract is less than \bar{c} . From (78), that will be the case whenever

$$k_2^* > (1 - \gamma) \delta g(\bar{c}) - \bar{c}, \quad (82)$$

or, equivalently,

$$k_1^* < k_2^* + [\beta g(c^*) - c^*] - (1 - \gamma) \delta g(\bar{c}) + \bar{c}.$$

When this condition is satisfied, it is certainly more profitable for firms to use training contracts even though it is efficient to train all workers.

References

- [1] Acemoglu, Daron. "Credit Constraints, Investment Externalities and Growth." In Booth and Snower (1996), 43–62.
- [2] Acemoglu, Daron. "Training and Innovation in an Imperfect Labour Market." *Review of Economic Studies*, 64 (July 1997), 445–464.
- [3] Acemoglu, Daron and Pischke, Jörn-Steffen. "Why Do Firms Train? Theory and Evidence." *Quarterly Journal of Economics*, 113 (February 1998), 79–119.
- [4] Acemoglu, Daron and Pischke, Jörn-Steffen. "Beyond Becker: Training in Imperfect Labour Markets." *Economic Journal*, 106 (February 1999a), F112–F142.
- [5] Acemoglu, Daron and Pischke, Jörn-Steffen. "The Structure of Wages and Investment in General Training." *Journal of Political Economy*, 107 (June 1999b), 539–572.
- [6] Becker, Gary. *Human Capital*. New York: Columbia University Press (1964).
- [7] Bindoff, S. T. *Tudor England*. Harmondsworth: Penguin (1950).
- [8] Booth, Alison L. and Chatterji, Monojit. *Training and Contracts*. Working Paper 95-15, Research Centre on Micro-Social Change, University of Essex (1995).
- [9] Booth, Alison L. and Chatterji, Monojit. "Unions and Efficient Training." *Economic Journal*, 108 (March 1998), 328–343.
- [10] Booth, Alison L. and Snower, Dennis J. *Acquiring Skills*. Cambridge: Cambridge University Press (1996).
- [11] Burdett, Kenneth and Smith, Eric. "Education and Matching Externalities." In Booth and Snower (1996), 65-80.
- [12] Cantor, Richard. "Firm Specific Training and Contract Length." *Economica*, 57 (February 1990), 1–14.
- [13] Chang, Chun and Wang, Yijiang. "Human Capital Investment under Asymmetric Information: The Pigovian Conjecture Revisited." *Journal of Labor Economics*, 14 (July 1996), 505–519.
- [14] Cunningham, W. and McArthur, E. A. *Outlines of English Industrial History*. Cambridge: Cambridge University Press (1920).
- [15] Dolton, Peter J. "The Economics of Youth Training in Britain." *Economic Journal*, 103 (September 1993), 1261–1278.
- [16] Earle, P. *The Making of the English Middle Class, 1660-1730*. London: Methuen (1989).
- [17] Elbaum, Bernard. "Why Apprenticeship Persisted in Britain But Not in the United States." *Journal of Economic History*, 49 (June 1989), 337–349.

- [18] Finegold, D. and Soskice, D. "The Failure of British Training: Analysis and Prescription." *Oxford Review of Economic Policy*, 4(3) (Autumn 1988), 21–53.
- [19] Greenwald, Bruce C. "Adverse Selection in the Labour Market." *Review of Economic Studies*, 53(3) (July 1986), 325–347.
- [20] Hamilton, Gillian. "Enforcement in Apprenticeship Contracts: Were Runaways a Serious Problem? Evidence from Montreal." *Journal of Economic History*, 55(3) (September 1995), 551–574.
- [21] Hamilton, Gillian. "The Market for Montreal Apprentices: Contract Length and Information." *Explorations in Economic History*, 33 (1996), 496–523.
- [22] Harhoff, Dietmar and Kane, Thomas J. "Financing Apprenticeship Training: Evidence from Germany." In N. Stacey (ed.). *School-to-Work: What Does Research Say about It?*, Washington, D.C.: Office of Educational Research and Improvement (ED) (1995).
- [23] Hermalin, Benjamin. "Adverse Selection, Short-Term Contracting, and the Underprovision of On-the-Job Training." Working Paper No. 90-139, Department of Economics, University of California at Berkeley (February 1990).
- [24] Kahn, Charles and Huberman, Gur. "Two-Sided Uncertainty and "Up-or-Out" Contracts." *Journal of Labor Economics*, 6 (October 1988), 423–444.
- [25] Katz, Eliakim and Ziderman, Adrian. "Investment in General Training: The Role of Information and Labour Mobility." *Economic Journal*, 100 (December 1990), 1147–1158.
- [26] Laffont, Jean-Jacques and Tirole, Jean. *A Theory of Incentives in Procurement and Regulation*, Cambridge, MA: MIT Press (1993).
- [27] Lee, D., Marsden, Dennis, Rickman, Penny and Duncombe, Jean. *Scheming for Youth: A Study of YTS in the Enterprise Culture*. Buckingham: Open University Press (1990).
- [28] Lees, Dennis and Chiplin, Brian "The Economics of Industrial Training." *Lloyds Bank Review*, 96 (1970), 29–41.
- [29] Loewenstein, Mark A. and Spletzer, James R. "Dividing the Costs and Returns to General Training." *Journal of Labor Economics*, 16 (January 1998), 142–171.
- [30] Lynch, Lisa M. "Private-Sector Training and the Earnings of Young Workers." *American Economic Review*, 82 (March 1992), 299–312.
- [31] Lynch, Lisa M. "The Economics of Youth Training in the U.S." *Economic Journal*, 103 (September 1993), 1292–1302.
- [32] Lynch, Lisa M. (ed.). *Training and the Private Sector: International Comparisons*, Chicago: University of Chicago Press (1994).
- [33] Pirenne, Henri. *Economic and Social History of Medieval Europe*. London: Kegan Paul, Trench, Trubner (1936).

- [34] Prendergast, Canice. "The Role of Promotion in Inducing Specific Human Capital Acquisition." *Quarterly Journal of Economics*, 108 (May 1993), 523-534.
- [35] Schlicht, Ekkehart. "Endogenous On-the-Job Training with Moral Hazard." *Labour Economics*, 3 (1996), 81-92.
- [36] Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Bell (1887).
- [37] Soskice, David. "Reconciling Markets and Institutions: The German Apprenticeship System." In Lynch (1994), 25-60.
- [38] Steedman, Hilary. "The Economics of Youth Training in Germany." *Economic Journal*, 103 (September 1993), 1279-1291.
- [39] Stevens, Margaret. "Labour Contracts and Efficiency in On-the-Job Training." *Economic Journal*, 104 (March 1994a), 408-419.
- [40] Stevens, Margaret. "An Investment Model for the Supply of Training by Employers." *Economic Journal*, 104 (May 1994b), 556-570.
- [41] Stevens, Margaret. "Transferable Training and Poaching Externalities." In Booth and Snower (1996), 21-40.
- [42] U.S. Department of Labor. *Report of the Commission on Workplace Quality and Labor Market Efficiency*. Washington: US GPO (1989).
- [43] Waldman, Michael. "'Up-or-Out' Contracts: A Signaling Perspective." *Journal of Labor Economics*, 8 (April 1990), 230-250.
- [44] Williamson, Oliver E. *The Economic Institutions of Capitalism*. New York: The Free Press (1985).