

The Review of Economic Studies Ltd.

Training and Innovation in an Imperfect Labour Market

Author(s): Daron Acemoglu

Source: *The Review of Economic Studies*, Vol. 64, No. 3 (Jul., 1997), pp. 445-464

Published by: The Review of Economic Studies Ltd.

Stable URL: <http://www.jstor.org/stable/2971723>

Accessed: 11/02/2009 10:38

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=resl>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



The Review of Economic Studies Ltd. is collaborating with JSTOR to digitize, preserve and extend access to *The Review of Economic Studies*.

<http://www.jstor.org>

Training and Innovation in an Imperfect Labour Market

DARON ACEMOGLU
Massachusetts Institute of Technology

First version received June 1995; final version accepted January 1997 (Eds.)

This paper shows that in a frictional labour market part of the productivity gains from general training will be captured by future employers. As a result, investments in general skills will be suboptimally low, and contrary to the standard theory, part of the costs may be borne by the employers. The paper also demonstrates that the interaction between innovation and training leads to an amplification of this inefficiency and to a multiplicity of equilibria. Workers are more willing to invest in their skills by accepting lower wages today if they expect more firms to innovate and pay them higher wages in the future. Similarly, firms are more willing to innovate when they expect the quality of the future workforce to be higher, thus when workers invest more in their skills.

1. INTRODUCTION

In modern economies, a large portion of human capital investments takes place within firms in the form of training. Most economic analyses of training are based on the paradigm introduced by Becker (1964) which suggests that a worker should pay for any general training which allows him/her to use the new skills when employed by other firms. Inefficiency occurs mainly because workers may be unable to pay for their training, and also unable to commit to not quitting their firm after employer sponsored training. Yet this source of inefficiency does not seem particularly relevant for many instances of training in the real world. For example, in Germany which has the most developed system of privately funded training, firms bear a large part of the cost of training while apprentices are often paid attractive wages (see for instance Harhoff and Kane (1994)) and also in the U.S. firms often pay for the vocational training courses of their employees (see Bishop (1991) or Acemoglu and Pischke (1996) for some numbers).

This paper argues that the focus on credit market imperfections ignores an important source of market failure in training. Although workers may be willing to pay the present-discounted value of their increased future earnings to a firm which offers training,¹ this present discounted value is only equal to the social value of training in a frictionless labour market. As soon as the hypothetical world of the Walrasian auctioneer is abandoned and labour markets characterized by costly mobility and search are considered, workers will not receive their full marginal product in future jobs. Because employer rents do not feature in workers' calculations, underinvestment in training will result. In other words, in an imperfect labour market future employers of a worker will also benefit from his skills. This is an externality that the decentralized market will not be able to internalize. It is important to contrast this mechanism to the one emphasized by the previous literature.

1. Throughout the paper, I will be talking of a firm "offering" training with the understanding that this is mostly on-the-job training, thus a worker cannot buy the same training in a school. Since the model has perfectly transferable utility, the addition of such an option would not change the results.

In this paper the externality is between the worker and his *future* employer whereas previous literature has concentrated on the inefficiencies in investment, due to incomplete contracts and credit market imperfections, between the worker and his current employer (see Becker (1964), pp. 93–95, Grout (1984)). These inefficiencies often have simple solutions as soon as the extreme incompleteness of contracts is relaxed (for instance in the form of exit penalties for workers who quit their firms). In contrast, contractual arrangements to deal with the externality proposed here are not easy to develop; the worker cannot contract with his future employer because at the investment stage it is unknown who this future employer will be. This is not only a more appealing source of inefficiency than the one emphasized by the previous literature, but also I will argue that it offers possible explanations to a number of otherwise puzzling observations, such as the willingness of employers to bear part of the costs of general training, and the absence of significant wage increases after such training is completed.

Training is most essential when new technologies are adopted, or in the process of a radical change of environment, for example, the shift from low- to high-skill jobs taking place in most OECD countries today. In support of this view, survey evidence suggests that the availability of appropriate skills is a key determinant of innovation and technology adoption decisions (e.g. see Northcott and Walling (1988), Northcott and Vickrey (1993)), and the efficient adoption of new technologies by Japanese firms is often attributed to their effective training strategies (e.g. Hashimoto (1991)). My analysis will establish not only that innovation decisions will be distorted due to labour market imperfections, but also that the interaction of innovation and training decisions leads to an amplification of inefficiencies and to a multiplicity of equilibria. Intuitively, skills are more valuable to a firm which has the new technology, thus the expected future wages of a worker depend on how many of his potential future employers have adopted the innovation. Similarly, expected profits depend on how skilled the future workforce is. As a result, when a greater number of firms adopt the innovation, workers expect higher wages from training and invest more in skills, and the profits from the new technology are higher. Therefore, the profitability of innovation and of training will increase with the *thickness* of the market for trained labour. An often-made claim is that capital investments in less developed economies are limited because the workforce is not sufficiently skilled (for instance this is suggested in Rosenstein–Rodan’s famous 1943 article pp. 204–205). My analysis formalizes this claim and shows that the workforce may be unskilled because of lack of capital investments. Moreover, I also demonstrate that this result crucially relies on labour market frictions.

This paper is related to a number of contributions in the literature. First, a number of recent papers have analysed the physical capital investments and schooling decisions in labour markets characterized by search, e.g. Acemoglu (1996*a*), Burdett and Smith (1996), Laing, Palivos and Wang (1995), Robinson (1995) and Saint-Paul (1996). For instance, Acemoglu (1996*a*) shows how the return to a worker can be increasing in the skill level of competing workers. Laing, Palivos and Wang (1995) demonstrate how education choices interact with wage determination and affect the growth rate in a search economy. In their paper, a more efficient matching technology encourages workers to invest more in education, and increases the growth rate. However, these papers neither deal with training choices nor identify the externality proposed in this paper. It is therefore not possible to understand why firms pay for training and whether training decisions will be efficient. Also, the interactions between adoption of new technologies and skills of workers, which appear to be important in practice, are not treated in the previous literature. Acemoglu and Pischke (1996) develop a model where workers do not pay for their general

training, but the key ingredient is the superior information that the initial employer has over future employers regarding the ability of their employees.

Another closely related literature concerns multiple equilibria in macroeconomics. This literature has identified two channels for agglomeration of activity and multiplicity of equilibria; the first is technological externalities, as in Durlauf (1993), and the second is aggregate demand externalities as in Kiyotaki (1988) or Murphy, Shleifer and Vishny (1989). This paper proposes a new source of multiplicity in macro-models arising from labour market imperfections. In my model there are neither technological externalities nor aggregate demand spillovers, but multiplicity of equilibria is likely to exist because skills are more valuable to workers when a greater number of potential employers with the new technology are hiring. Diamond's famous (1982) contribution on search is also motivated by a labour market analogy, but the source of multiplicity is increasing return (externalities) in the matching technology. There are other examples of multiplicity in frictional labour markets but the mechanisms are rather different. In Laing, Palivos and Wang (1995) and Robinson (1995) more schooling leads to higher firm profits; this induces further entry by firms, increasing the likelihood of employment and the incentives to invest in skills for workers. In contrast, the source of multiplicity in this paper is the fact that skills are more valuable to firms with the new technologies combined with the externality identified between the worker and his *future* employers.

The plan of the paper is as follows. Section 2 presents the environment, derives the efficient allocation and characterizes the Walrasian equilibrium. Section 3 contains the main results of this paper. First, I abstract from innovation decisions and show how underinvestment in training occurs, then I illustrate how multiple equilibria can emerge from the interaction of training and innovation. Section 4 includes some extensions. The appendix contains a more detailed analysis of the equilibrium of the economy in the absence of labour market frictions.

2. THE MODEL AND COMPETITIVE EQUILIBRIUM

2.1. *Description of the environment*

Consider an economy consisting of a continuum of risk-neutral workers and risk-neutral firms. Each group has mass 1. All agents have discount rate equal to r and enjoy the consumption of the only good of this economy. The economy lasts for T periods. I will start with the case of $T=2$ and will later consider $T=\infty$ to give an idea of the possible magnitude of the effects discussed here and also to illustrate the possibility of delay. Each firm has access to a Leontieff production function which requires one worker and produces per period output equal to y . A worker or firm on its own produces nothing.

Output of a pair can be increased by investment. There are two types of investments which can only be undertaken in period 1 and affect productivity in periods $t \geq 2$. The first is an investment in new technology (or innovation). At cost δ the firm can acquire a new machine. The important assumption here is that it is necessary for the property rights over the machine to be vested in the firm. If the employment relation between the firm and the worker ends, the machine will stay with the firm. The second type of investment is in the *general* human capital of the worker. The worker can acquire τ units of general human capital, but this reduces the output of the firm in the first period by $c(\tau)$. Whether the firm or the worker incurs this cost is immaterial since throughout the paper I assume that there are no credit constraints or equivalently that utility is transferable. Human capital is general in the sense that the worker can use his skills with any firm. Also

Assumption 1. $c(\cdot)$ is differentiable, strictly increasing and convex and $c(0) = 0$,
and $\exists \bar{\tau}$ s.t. $\lim_{\tau \rightarrow \bar{\tau}} c(\tau) = \infty$.

Therefore, the level of training is chosen from a set $[0, \bar{\tau}]$ and the cost function is strictly convex. Let $\gamma_j = 0$ denote that firm j does not have the new technology and $\gamma_j = 1$ that it has the new technology, and τ_i is the training level of worker i . The productivity of worker i and firm j in all periods $t \geq 2$ is assumed to be equal to $y + \alpha(\gamma_j, \tau_i)$.

Assumption 2. $\alpha(0, \tau) = \alpha_0 \tau$ and $\alpha(1, \tau) = \alpha_1 \tau$, $\forall \tau$, and $\alpha_0 \leq \alpha_1$.

This assumption implies that training and investment in new technology are complements. This is important for the multiplicity result, but also very plausible. A sizeable empirical literature starting with Griliches (1969) establishes the complementarities between physical and human capital, and Bartel and Lichtenberg (1989) show that firms which adopt new technologies are more likely to train their workers.

There is a final feature of the technology which needs to be specified. At the end of every period that worker i and firm j are together, there is a probability $s \in (0, 1)$, that the pair receives an adverse match-specific shock which reduces their output to 0 in all future periods. After such a shock, both parties can try to find new partners for production.

It is important to reiterate that since the output of a pair does not depend on the decisions of other agents in this economy, there are no technological externalities. Further, since utility is perfectly *transferable* across all agents, there are no credit market constraints nor aggregate demand externalities.

2.2. The Pareto optimal allocation

Since there are no social costs of allocating workers to different firms in the second period (i.e. s does not matter), the efficient allocation can be determined by maximizing the joint surplus of the partnership with respect to γ and τ , which is

$$\max_{\gamma, \tau} \alpha(\gamma, \tau) - (1+r)(c(\tau) - \gamma\delta). \quad (1)$$

This problem will have a solution with either $\gamma = 0$ or 1. It is clear that if $\gamma = 0$, the optimal level of training is given by τ^l such that

$$\alpha_0 = (1+r)c'(\tau^l).$$

In contrast, when $\gamma = 1$, the efficient level of training is τ^h such that

$$\alpha_1 = (1+r)c'(\tau^h).$$

Given the strict convexity of $c(\cdot)$, both τ^l and τ^h are uniquely defined.

Assumption 3.

- (i) $\alpha_1 \tau^h - (1+r)(c(\tau^h) + \delta) > \alpha_0 \tau^l - (1+r)c(\tau^l)$.
- (ii) $\alpha_1 \tau^l - (1+r)(c(\tau^l) + \delta) < \alpha_0 \tau^l - (1+r)c(\tau^l)$.

Part (i) of the assumption ensures that (1) has a unique solution where $\gamma = 1$ and $\tau = \tau^h$, thus investment in the new technology and training are socially efficient. If Assumption

3 did not hold, then $\gamma = 0$ and $\tau = \tau'$ would give the Pareto optimal allocation. The second part of the assumption restricts attention to the area of the parameter space where the effects emphasized in this paper can be seen most clearly: it ensures that when all workers choose τ' , investment in the new technology is no longer desirable.

2.3. *Equilibrium without frictions*

I now characterize the equilibrium of this economy in the absence of labour market frictions. In this subsection, I will only state the existence of a unique and Pareto optimal Walrasian allocation, which is not surprising in view of the second welfare theorem. The Walrasian allocation is not the only possible benchmark equilibrium in the frictionless case, and it is in fact quite a stringent equilibrium concept as it requires all trades to take place at time $t = 1$. An alternative is to characterize the subgame perfect equilibria of the game in which firms and workers trade in each period and there are no frictions nor costs of changing partners. This different approach is taken in the Appendix and will yield the same result.

To define the equilibrium concept more formally, let $w(\tau)$ be the price at which the labour services of a worker with training τ is traded at time $t = 2$, and $v(\tau)$ be the price at which a firm can hire the labour services of a worker at time $t = 1$ if it promises to provide an amount of training τ to this worker (note that there is no enforceability problem related to this promise). Also denote the excess demand for the services of workers with training τ at time $t = 2$ by $e_2(\tau)$, and the excess demand from firms offering training τ at time $t = 1$ by $e_1(\tau)$. A Walrasian equilibrium is a set of wage functions $w(\tau)$ and $v(\tau)$ such that $w: [0, \bar{\tau}] \rightarrow \mathbb{R}^+$ and $v: [0, \bar{\tau}] \rightarrow \mathbb{R}$ (thus v can take negative values—the worker paying to receive the training), such that $\forall t, \tau, e_t(\tau) \leq 0$, and if $w(\tau) > 0$, then $e_2(\tau) = 0$ and if $v(\tau) > 0$, then $e_1(\tau) = 0$.

Proposition 1. *There exists a unique Walrasian equilibrium allocation in which all firms adopt the new technology and all workers choose $\tau = \tau^h$.*

The proof follows immediately from the discussion in the Appendix. This proposition states that the unique Walrasian equilibrium is Pareto optimal. Therefore, in the rest of the paper, I will take the Pareto efficient allocation as the unique benchmark equilibrium in the absence of labour market frictions. It is also straightforward to see that in this unique equilibrium, the worker is bearing the full cost of training, in the sense that $v(\tau) = v(0) - c(\tau)$. In other words, in the first period, the worker takes a wage cut equal to the cost of training he is receiving. This is an important, and well known (see Becker (1964)), feature of the frictionless economy.

3. EQUILIBRIUM IN AN IMPERFECT LABOUR MARKET

3.1. *Description*

All the technological features are the same as above but the market mechanism is different. In particular, there is no spot labour market with an auctioneer calling out wage functions to regulate trade. Instead, in each period, workers and firms looking for a new partner have to engage in costly search as in the models of Diamond and Maskin (1979), Diamond (1982), Mortensen (1982) and Pissarides (1990). In particular, agents will be matched one to another randomly, and all parties will have a partner, thus there will be no unemployment. The assumption that every agent finds a partner is in order to demonstrate that the

difference between the competitive and non-competitive economies is not due to the presence of unemployment (see Section 4.3 for unemployment). There is however no guarantee that the firm with the investment in the new technology will be matched with the worker who has more training: thus the matching technology is random (as is the norm in this type of models, see for instance, Sattinger (1995), Burdett and Coles (1995), or Acemoglu (1996b)). This is an important difference between the competitive and frictional environments since in the absence of frictions, the high-training workers will always produce with the firms which possess the new technology (see Lemma 1 in the Appendix for more details). This aspect of the technology will be discussed in more detail later. Finally, to simplify the analysis I assume that after the initial random match between a firm and a worker, breaking the match and looking for a new partner within the same period is excessively costly, thus workers and firms who are matched together will find it profitable to reach agreement.²

How are wages determined? Since there is no spot labour market, wages cannot be equal to the opportunity cost (outside option) of both sides. This implies that there will be some *rent-sharing*. The most common way of dealing with this is bargaining, and any bargaining rule will yield the same results. Here I adopt the bargaining solution suggested by Shaked and Sutton (1984) [see Acemoglu (1996a) Appendix A for details and a game-theoretic derivation in the context of search models, see also Binmore, Rubinstein and Wolinsky (1985)]. According to this bargaining rule, as long as both parties are getting more than their outside options, the gross surplus of the partnership is shared. In the current context this simply implies that the worker will get a proportion β of the total surplus and the firm obtains the remaining $1 - \beta$ proportion. Also, in order to emphasize the novel source of inefficiency in this paper, rather than simply force the firm and the worker to bargain over current output, I assume that a firm and a worker who are together can write a *completely enforceable and binding long-term contract* to determine the future divisions of surplus. More explicitly, worker i and firm j , if together at time $t=1$, can write a contract that specifies future contingent transfers between themselves and therefore will share the present discounted value of the total surplus of the partnership.

3.2. Underinvestment in training

This subsection will demonstrate inefficiency in training investments due to labour market imperfections. We will also find that firms could now be willing to bear part of the costs of general skills as we observe in practice. To focus on the main point, I assume that $\alpha_0 = \alpha_1$, and $\delta = 0$, thus the decision to invest in the new technology is irrelevant. In this case the Pareto optimal (and Walrasian equilibrium) amount of training is given by $\tau^h = \tau^l$ (since $\alpha_1 = \alpha_0$).

2. Formally, a firm and a worker incur a cost ζ when they break a match within a period and find a new partner (for instance, the value of foregone production in the process of search). The limit point with $\zeta = 0$ is the competitive equilibria analysed in the Appendix. Here, I am assuming that ζ is large enough that no break-up ever occurs. This assumption is of no major importance. Appendix A of Acemoglu (1996a) proves, in a similar setting, that even for very small costs of breaking-up, the same equilibrium will be obtained. The same proof can be applied in this setting quite easily, but since it is of tangential interest for this paper, I will only give the intuition. All the equilibria of interest will be symmetric in the sense that all workers will have the same level of training and all firms will have the same technology. Thus, by breaking up, the firm will find itself in exactly the same bargaining situation with another worker, and in the subgame perfect equilibrium, it will get the same share of the surplus (see Shaked and Sutton (1984)). This reasoning also applies to workers. Therefore, for all positive values of ζ , the conclusions of the analysis of large ζ will apply. Only at the limit point of no frictions, $\zeta = 0$, the results of the previous section and the Appendix will be valid.

To characterize the equilibrium, start with the problem of worker i and firm j' who meet in period $t=2$ after being separated from their respective partners. Given the assumptions, they will simply split the overall surplus which is equal to $y + \alpha_0\tau_i$, thus the worker gets a wage $\hat{w}_i = \beta(y + \alpha_0\tau_i)$. I can now analyse the problem of worker i and firm j at $t=1$. Recall that this pair can write a long-term contract which specifies; (i) a first-period wage; (ii) training level; (iii) a second-period wage if the firm and the worker are together in the second period; (iv) when to terminate the relation; (v) a transfer from one party to the other if there is a termination.

If the adverse shock causes the productivity of the pair to diminish to zero, it is mutually beneficial to terminate the relation (by terminating both parties will get a positive return whereas without termination, they will both get zero), and the worker pays the firm the agreed transfer f . If there is no adverse shock, worker i and firm j produce together in the second period and the firm pays the worker the contractually determined wage w . Let us denote the training level by τ , the first period wage by v , and the distribution function for training among workers by $q(\tilde{\tau})$. Since the firm and the worker can write a long-term contract and utility is perfectly transferable, τ , v , w and f will simply be chosen to maximize the total surplus of the relationship. This total surplus is given as

$$TS = \frac{(1-s)[y + \alpha_0\tau] + s[\beta(y + \alpha_0\tau) + (1-\beta)(y + \alpha_0 \int \tilde{\tau} dq(\tilde{\tau}))]}{1+r} - c(\tau). \tag{2}$$

Let us now carefully review the terms that make up (2). The last term is the cost of training that the relationship incurs immediately. All benefits accrue in the second period, hence the discounting. With probability $(1-s)$, there is no adverse shock and the relationship continues. In this case, irrespective of the value of the second period wage w , the total surplus is $y + \alpha_0\tau$, thus w only determines the distribution of the surplus. The more involved case occurs with probability s , when there is a separation. In this case the worker pays f to the firm. Nevertheless, f does not feature in expression (2) since it is a pure transfer.³ This is similarly the reason why w and v are not in (2); they do not influence the marginal incentive to invest in training. After the separation, the firm and the worker find new partners and bargain. The worker obtains a proportion β of the surplus with his new firm which gives $\hat{w} = \beta(y + \alpha_0\tau)$ as his second-period wage. Similarly, the firm obtains a proportion $1-\beta$ of the output that the new worker produces, thus its return is $(1-\beta)(y + \alpha_0\tilde{\tau})$. (2) is obtained by taking expectations over $\tilde{\tau}$ and noting that the subsample of workers who are looking for a match in the second period is randomly drawn from the population of workers since all pairs face the same probability of separation. The following proposition is immediate⁴:

Proposition 2. *There is a unique equilibrium in which all pairs choose a level of training $\hat{\tau}$ in the first period such that $\alpha_0[(1-s) + s\beta] = (1+r)c'(\hat{\tau})$ and thus $\hat{\tau} < \tau^h = \tau^l$. In this unique equilibrium, $(1+r)v + (1-s)w - sf = \beta[(1+r)y + y + \alpha_0\hat{\tau}]$.*

3. The reason I have included f in the discussion so far is that the presence of such an exit fee would prevent most of the inefficiencies in training in models with competitive labour markets and credit constraints, thus the presence of this variable emphasizes that this type of incompleteness of contracts is not related to the mechanism proposed here.

4. Note that the separation decision is assumed to be specified and committed to. In the absence of this assumption none of the results would change but w and f need to be restricted to $w = \beta(y + \alpha_0\hat{\tau}) - f$ because with a lower second period wage, the worker would prefer to leave the firm and find a new employer, and similarly for a higher wage, knowing that all other workers also have training $\hat{\tau}$, the firm would prefer to separate.

Proof. Differentiating (2) with respect to τ gives a unique solution $\tau = \hat{\tau}$, thus all workers choose this level of training. Comparison with the F.O.C. of the planner's problem immediately implies that $\hat{\tau} < \tau' = \tau^h$. Finally, given $\hat{\tau}$, the total surplus to be divided is $y + (y + \alpha_0 \hat{\tau} / (1 + t))$ and the worker gets a share β of this which can be made up of different linear combinations of v , w and f . ||

There is a unique equilibrium with underinvestment in training. The reason is intuitive. In the competitive equilibrium, in the second period the worker obtained 100% of the increase in productivity due to training, and was therefore willing to pay the cost. In contrast, here with probability s an *unknown third-party*, the future employer, is getting a proportion $(1 - \beta)$ of the benefit. Underinvestment in training results because the rents accruing to this third party do not feature in the calculations of the worker and his current employer.

Although the firm and the worker can write complicated and binding contracts, the externality is not internalized. The positive externality is onto the future employer, therefore the crucial variable that the firm and the worker would like to contract upon is \hat{w} , the wage in the second period *after* the worker separates from the current firm and finds a new employer. However, at the time of training it is not known who this future employer will be, thus contracting with this firm is not possible. This incompleteness of contracts between the worker and his future employers is therefore crucial for the underinvestment result.

This discussion also implies that there are two types of contracts, not allowed in the analysis, which would help with the inefficiency. The first, *contract 1*, would involve the firm and the worker writing in their contract that if in the second period the worker has a new employer, this new employer has to pay a certain amount, $P(\tau)$, to the initial firm, otherwise the worker is not allowed to work. The second, *contract 2*, works similarly, but requires the new employer of the worker to pay a certain wage to the worker. For instance, conditional on a separation in the second period, if $\hat{w} < x$, then $\hat{f} = F$ and if $\hat{w} \geq x$, then $\hat{f} = 0$ where \hat{f} is the punishment imposed on the worker for breaking the terms of the contract (say as a payment to his initial employer). It is straightforward to obtain the following corollary to Proposition 2.

Corollary 1. *If contract 1 can be written and enforced, then $P(\tau) = \alpha_0 \tau$ would implement the efficient training τ^h . If contract 2 can be written and enforced, then the efficient level of training, τ^h , can be implemented with $x = \alpha_0 \tau$, and F large enough.*

The proof of this corollary is immediate from the discussion and is thus omitted. Both contracts work in the same way: they impose an obligation on the future employer by forcing it to make a higher payment. This is of course intuitive since as noted above, training creates a positive externality on the future employer, and these contracts ensure that this firm contributes to the costs of training. In the first case, the payment imposed on the future employer is a transfer fee to the initial employer, and in the second it is a sufficiently high wage payment to the worker. Since complete contracts between the worker and his initial employer are possible, whether the initial firm or the worker gets the future payments is of no consequence, hence both contracts essentially achieve the same goal.

However, both types of contracts are very difficult to implement precisely because they impose obligations on a party who is not part of the contract, and are thus not legally enforceable. In fact, contracts which enable a worker to unilaterally determine his future wages (by committing not to work for less) would certainly distort a number of important

economic decisions (e.g. firm entry and investments), and would be very hard to enforce. Also in a more general model, contracts of this type will often hurt the worker by reducing his future employment opportunities. There will also be some more specific problems with these contracts, for instance, with *contract 2* if the worker and the new firm can distort the observation of \hat{w} , the bargaining between them will lead to a wage less than x , but they will report the wage to be x . When this possibility is present, anticipating that the equilibrium wage \hat{w} will be less than x , training investments in the first period will still be suboptimal.

It is also interesting to note that an example of *contract 1* was actually used in the market for European football players. This contract which required a new team to pay a transfer fee to the player's previous club has been recently challenged and declared unlawful in European courts precisely on such grounds (the Bosman Case, September, 1995). Overall, it is safe to presume that such contracts are not possible and that there will therefore be underinvestment in training. Nonetheless, despite the implementation difficulties, the fact that such contracts were tried suggests that the externality this paper emphasizes is important. Interestingly, if the market failure emphasized by the previous literature (i.e. externalities between the worker and his initial employer) was substantial, we would expect the transfer fee to be from the player to his previous club. Although star soccer players are far from being liquidity constrained, such contracts are not observed.

The emphasis on the inability to contract with future employers does not however imply that the results are driven by incomplete contracts rather than frictional labour markets. Both the need to write such a contract and the inability to do so are intimately related to the frictions in the labour market: in the perfectly competitive economy of Section 2.5, because all workers are always paid their marginal product, there was no need to write a contract with the future employers. Also, it is precisely the decentralized nature of the labour market which makes it impossible for the worker to sign a contract at $t=1$ with all possible future employers (i.e. a grand contract between all workers and all firms). Therefore, the key results of this paper are driven by labour market frictions, but they also rely on the impossibility to write contracts with future employers.

Another important implication of this approach is that v , w and f are not individually determined; the only equilibrium requirement is that the overall share of the worker be equal to a proportion β of the total surplus. Therefore, it is not possible to infer the productivity of training from the wage-profile of the worker. This observation explains a number of otherwise puzzling findings. For instance, Pischke (1996) looks at the further training offered by German firms. Because this is a voluntary activity on both sides, both the worker and the firm must benefit. Yet, although workers report that the skills they acquire are general, there is no increase in the wage after training. My model suggests that this puzzling observation may be due to the fact that the worker is paid for his future productivity increase during training (i.e. high v and low w), and more importantly, in contrast to the economy with perfectly competitive labour markets, he cannot costlessly change employer and obtain his full marginal product after training; in other words, the firm is able to capture some of the increase in productivity due to the additional skills. Therefore, in contrast to the competitive benchmark as analysed by Becker (1964), the firm will often have an incentive to bear part of the training costs, because due to hold-up by future employers, the worker will only have limited mobility. Therefore, the model offered in this paper is capable of explaining the prevalence of firm sponsored general training (see Harhoff and Kane (1994) for German, Jones (1986) for British, and Bishop (1991) for U.S. evidence on this).

Finally, the model also predicts an inverse relation between s (turnover) and train- In particular, when $s=0$, there are no external effects on future employers, and the

equilibrium is efficient. The greater is s , the larger is the likelihood that the skills of the worker will be used by some future employer, and the more serious is the inefficiency. This negative correlation between training and turnover is often found in the data (see for instance OECD (1994)) and is argued to be related to market failure in training in Anglo-Saxon economies (e.g. Blinder and Krueger (1991)). Although the standard Beckerian model of training would predict an inverse relation between turnover and *firm-specific* human capital, a large portion of the investments that take place in German and Japanese firms appears to be general (see for instance Tan (1990), Harhoff and Kane (1994), Acemoglu and Pischke (1996)). Moreover, it is socially optimal to have an inverse relation between firm-specific training and turnover. Therefore, existing models neither explain the high levels of general training in low turnover markets nor substantiate the claim that low training induced by high turnover is a market failure. My model accounts for the negative correlation and suggests why this may be indicative of inefficiencies in the amount of general training.

3.3. *New technologies and multiple equilibria*

In this subsection I reintroduce the technology choice as in Section 2 and demonstrate that the interaction between innovation and training can lead to the existence of multiple equilibria. The presence of multiple equilibria can be interpreted to imply that similar economies may have very different outcomes due to small differences in environments; or that the inefficiency introduced in the last subsection is much amplified as a result of the new interactions; and/or that some less developed economies do not generate investments in new technologies and have highly unskilled labour forces as a result of a coordination failure.

The same ingredients as before lead us to write a programme similar to (2) to be maximized with respect to γ and τ (again w , v and f are just transfers)

$$\begin{aligned} \max_{\gamma, \tau} TS = & \frac{(1-s)(y + \alpha_0\tau + \gamma(\alpha_1 - \alpha_0)\tau)}{1+r} \\ & + \frac{s(\beta[y + (1-\phi)\alpha_0\tau + \phi\alpha_1\tau] + (1-\beta)[y + ((1-\gamma)\alpha_0 + \gamma\alpha_1) \int \tilde{\tau} dq(\tilde{\tau})])}{1+r} \\ & - [c(\tau) + \gamma\delta], \end{aligned} \quad (3)$$

where ϕ is the proportion of firms with the new technology, and since s hits pairs irrespective of their characteristics, ϕ is also the proportion of firms with the new technology in the subsample of firms looking for a new worker in period $t=2$. The rest of the expression (3) has exactly the same intuition as (2), the only new feature being that when $\gamma=1$, there is an increase in productivity equal to $(\alpha_1 - \alpha_0)\tilde{\tau}$ for a firm employing a worker with training $\tilde{\tau}$. Note that if the adverse productivity shock does not hit a pair, they are assumed not to separate (see next section).

The next result characterizes the equilibrium of this economy. First I introduce two conditions which will be useful in the statement of this proportion.

Condition 1.

$$\alpha_1 \hat{\tau}^h - (1+r)(\delta + c(\hat{\tau}^h)) > (1-s)\alpha_0 \bar{\tau}' + s\beta\alpha_1 \bar{\tau}' + s(1-\beta)\alpha_0 \hat{\tau}^h - (1+r)c(\bar{\tau}')$$

Condition 2.

$$\alpha_0 \hat{\tau}^l - (1+r)c(\hat{\tau}^l) > (1-s)\alpha_1 \bar{\tau}^h + s\beta\alpha_0 \bar{\tau}^h + s(1-\beta)\alpha_1 \hat{\tau}^l - (1+r)(\delta + c(\bar{\tau}^h));$$

where $\hat{\tau}^h$, $\bar{\tau}^h$, $\hat{\tau}^l$ and $\bar{\tau}^l$ are defined as: $(1-s)\alpha_1 + s\beta\alpha_1 = (1+r)c'(\hat{\tau}^h)$, $(1-s)\alpha_0 + s\beta\alpha_0 = (1+r)c'(\hat{\tau}^l)$, $(1-s)\alpha_0 + s\beta\alpha_1 = (1+r)c'(\bar{\tau}^l)$ and $(1-s)\alpha_1 + s\beta\alpha_0 = (1+r)c'(\bar{\tau}^h)$.

In words, $\hat{\tau}^h$ is the level of training a pair would choose when they decide to adopt the new technology and when they anticipate $\phi=1$ (i.e. all other firms are also adopting the new technology). The only difference between $\hat{\tau}^h$ and the efficient amount τ^h arises because of the underinvestment effect identified in the previous subsection; the firm and the worker recognize that with probability s , the worker will have a new employer who will capture a proportion $1-\beta$ of his additional productivity. In contrast, $\bar{\tau}^h$ is the level of training when they decide to adopt the new technology but anticipate $\phi=0$. Similarly $\hat{\tau}^l$ is the level of training when the pair decides not to adopt the new technology and anticipates $\phi=0$. And $\bar{\tau}^l$ is the level of training if the innovation is not adopted but $\phi=1$. It is straightforward to see that $\tau^h > \hat{\tau}^h > \bar{\tau}^h$ and $\tau^l > \hat{\tau}^l$ and $\bar{\tau}^l > \hat{\tau}^l$. Therefore, Condition 1 is more restrictive than Assumption 3(i), and although Assumption 3 still holds, Condition 1 may not be satisfied. Also Conditions 1 and 2 can hold together, but not that given Assumption 3, it is not possible that both conditions be violated at the same time. Then, we have:

Proposition 3. *There can be two types of pure strategy symmetric equilibria:*

A: *All pairs innovate ($\phi=1$) and all workers choose $\tau = \hat{\tau}^h$.*

B: *No pair innovates ($\phi=0$) and all workers choose $\tau = \hat{\tau}^l$.*

(i) *If Conditions 1 and 2 hold, then both A and B are equilibria and A Pareto dominates B.⁵*

(ii) *If Condition 1 holds and 2 does not, then A is the unique equilibrium,*

(iii) *If Condition 2 holds and 1 does not, then B is the unique equilibrium.*

Proof. Suppose $\phi=1$ and worker i and firm j decide to adopt the technology. Then the F.O.C. of (3) with respect to τ gives $\hat{\tau}^h$. When all other pairs expect $\phi=1$ and innovate they will also choose $\hat{\tau}^h$, thus $q(\tau)$ has all its mass at $\hat{\tau}^h$, therefore the return from adopting the innovation is given by the LHS of Condition 1. Now suppose that worker i and firm j deviate and do not innovate. Since $\phi=1$, the F.O.C. will give $(1-s)\alpha_0 + s\beta\alpha_1 = c'(\tau)$ which is $\bar{\tau}^l$. Because all other workers have training level $\hat{\tau}^h$, in case of separation the firm expects an additional $(1-\beta)\alpha_0 \hat{\tau}^h$, thus the profit to deviating is given by the RHS of Condition 1. If Condition 1 holds, then deviating is not profitable and there exists an equilibrium with $\phi=1$ and $\tau = \hat{\tau}^h$. Similarly, if $\phi=0$, a pair not innovating would choose $\hat{\tau}^l$ and anticipating that $\tau_i = \hat{\tau}^l$ for all workers, the expected surplus will be given by the LHS of Condition 2. If a pair deviates and decides to innovate, then they will choose a level of training $\bar{\tau}^h$, and the total surplus is the RHS of Condition 2. Thus when this condition holds, there is an equilibrium with $\phi=0$ and $\tau = \hat{\tau}^l$. If both conditions hold, both allocations are equilibria. Since Assumption 3(i) holds, high training is preferred and thus at $\hat{\tau}^l$ and $\gamma=1$, all agents are better off compared to an allocation with $\hat{\tau}^l$ and no innovation. ||

5. In this case, there also exists a mixed strategy equilibrium where firms are indifferent between $\gamma=1$ and $\gamma=0$ and all workers choose $\tau^m \in (\hat{\tau}^l, \hat{\tau}^h)$.

For the set of parameter values at which there is a unique Pareto optimal competitive equilibrium, the economy with costly search can have multiple Pareto ranked equilibria or can have a unique equilibrium with no innovation and low training. The failure to capture all the rents created by training which led to underinvestment in the previous subsection is still present, thus training and innovation are less profitable than in the competitive equilibrium. Moreover, the private benefits from training depend on the expected wage of the worker in the case of separation. From expression (3), this expected wage is $\beta[y + (1 - \phi)\alpha_0\tau + \phi\alpha_1\tau]$ and is increasing in ϕ . If a greater number of firms possess the new technology, it is more likely that the marginal product of the worker will be $\alpha_1\tau$ rather than $\alpha_0\tau$. Hence the larger is ϕ , the higher is the desired level of training. Now let us look at the problem facing the firm. The profitability of innovation will depend on expected profits in case of separation which is $(1 - \beta)[y + ((1 - \gamma)\alpha_0 + \gamma\alpha_1) \int \tilde{\tau}dq(\tilde{\tau})]$. Therefore, expected profits after a separation are increasing faster in γ when the expected skill level of the *future* workforce, $\int \tilde{\tau}dq(\tilde{\tau})$, is higher. Thus, firms are more willing to invest in new technologies when they expect future employees to be skilled. Further, when other firms are adopting the innovation, a pair that adopts the technology will choose a higher level of training, ϕ will be larger, average training will be higher, and innovation will be more profitable hence the *multiplicity of equilibria*. Therefore, in the range of parameter values where it is socially beneficial to innovate, there can now be two equilibria: in one all firms innovate and workers have a relatively high level of skill, $\hat{\tau}^h$. In the other, there is a coordination failure, firms do not innovate expecting their future recruits to only have a low level of skill, $\hat{\tau}^l$, and workers expecting the firms not to innovate choose $\hat{\tau}^l$. Note again that as s tends to zero, the inefficiencies disappear, and Conditions 1 and 2 become identical. Thus, without separations, there are no inefficiencies nor multiple equilibria because there are no interactions with future employers, and so the externality identified in this paper disappears.

An interesting question is whether the random matching assumption, that a high-training worker has exactly the same probability of matching with a firm with the new technology as a low-skill worker, is important for our results. It is first easy to see that the results of subsection 3.2 on inefficient training investments are not affected: these results were derived with all firms having the same technology, and were driven only by the impossibility to write contracts with future employers. In this subsection, in contrast, random matching plays an important role since there is potential heterogeneity on both sides. In particular, the intuition for the multiplicity of equilibria is that the *thickness* of the market for skilled labour, ϕ , matters: when only a few firms have the new technology, it is not worthwhile to acquire additional skills. This intuition does not hold with competitive markets because as Lemma 1 in the Appendix proves, irrespective of the value of ϕ , the same wage rule applies. However, the extreme random matching assumption is of no major consequence. This is because both Conditions 1 and 2 are for the cases in which all firms make the same decision. Therefore, irrespective of the matching technology (as long as we are not in a fully Walrasian market), these conditions will be valid.⁶ Thus even when the assumption of random matching is relaxed (e.g. high skill workers having a higher probability of matching with high technology firms, see Acemoglu (1997)), the results of this paper would not be affected at all. This again reiterates the main conclusion of the previous subsection that some degree of labour market frictions and the inability

6. Although Conditions 1 and 2 would not change, equation (3) would change, in particular, the term $\int \tilde{\tau}dq(\tilde{\tau})$ would be replaced with one which takes the differential matching probabilities into consideration.

to write contracts with future employers are the crucial ingredients for the results of this paper.

3.4. A simple example and some numbers

To obtain the intuition more clearly, and to help us assess the possible magnitude of the mechanism proposed here, consider the following simple example where τ takes the values 0 and 1 with $c(0)=0$ and $c(1)=c$. Thus if there is training and innovation, productivity goes up by α_1 and because skills and technology are strongly complementary, without either training or innovation there is no change in productivity. It is immediate to see that Assumption 3(i) now takes the form of $\alpha_1 > (1+r)(\delta+c)$. Let us write the expected increase in surplus due to innovation. This will be equal to $(1-s)\alpha_1 + s\beta\phi\alpha_1 + s(1-\beta)\alpha_1 \int \tilde{\tau}dq(\tilde{\tau}) - (1+r)(\delta+c)$. Since in this case a pair would only invest in training when they adopt the new technology, we have $\int \tilde{\tau}dq(\tilde{\tau}) = \phi$, and thus innovation and training are profitable if $[(1-s) + s\phi]\alpha_1 > (1+r)(\delta+c)$. By substituting $\phi=1$, we can see that Condition 1 is exactly the same as Assumption 3(i). Therefore, an equilibrium with high training and innovation always exists. Similarly, Condition 2 becomes $(1-s)\alpha_1 < (1+r)(\delta+c)$. In this case Conditions 1 and 2 are satisfied for $(1-s)\alpha_1 < (1+r)(\delta+c) < \alpha_1$, which is a non-empty set of parameter values. The under-training result of the previous subsection does not arise here and Assumption 3(i) and Condition 1 coincide because training decisions are binary, and as a result the loss of surplus to the future employer does not influence the marginal training decision.

In this simple example it is easier to see the exact reason for the multiplicity of equilibria. In the Walrasian economy a trained worker will always find a firm with the new technology which can make use of his skills, thus the marginal return to more training is always α_1 irrespective of the value of ϕ (see Lemma 1 in the Appendix). In contrast, in the economy with frictional labour markets, the likelihood of finding a firm with the new technology is ϕ , and the likelihood of finding the right worker for the firm is ϕ . Therefore, the *thickness* of the market for trained workers (and for firms with the new technology) plays a crucial role in the absence of the Walrasian auctioneer. This reasoning also establishes that for the claim that less developed economies may have limited investment because of a low-skill workforce to be true, labour market imperfections are necessary; this claim is not true in a competitive labour market.

How important is this interaction? Take $T = \infty$, and consider the example with only $\tau=0$ and $\tau=1$ discussed above. When all other firms adopt the innovation, the profit is equal to $\alpha_1/r - c - \delta$. Whereas when no other firm adopts the innovation, the return is equal to $(1-s)\alpha_1/(r+s) - c - \delta$. Thus, the expected return from innovation and training will be higher by a factor of $s(1+r)/(r+s)$ when all other firms invest compared to when they do not. Studies such as Leonard (1987) and Blanchard and Diamond (1990) find monthly separation rates as high as 4.5%. A recent study by Andersen and Meyer (1994) calculates a total quarterly separation rate of 23%. For my calculation it is more appropriate to concentrate on the permanent part of this rate which is 17%. Also in line with other evidence, Andersen and Meyer find that the likelihood of permanent separations fall with tenure. To arrive at a conservative estimate I use the lowest hazard rate that Andersen and Meyer calculate for workers who have been with their firm for more than 16 quarters. This gives a quarterly value of 5.7% for s . I take the quarterly real interest rate to be 1%. These numbers imply that the return from new technology will be higher by roughly 86% when all other firms adopt the new technology and train their workers compared to the

case in which no other firm invests. This suggests that the mechanism proposed in this paper can be a very important determinant of innovation and training decisions.

4. EXTENSIONS

4.1. *Delay and inefficiencies*

The assumption so far has been that all investments have to be made during period $t=1$. If a firm does not adopt the innovation at this date, it can never adopt it in the future. This is a strong assumption since most new technologies can be adopted any time after they become available. To show how the results change, I use the simple example introduced at the end of Section 3 where $\tau=0$ or $\tau=1$ and take the horizon to be infinite, i.e. $T=\infty$. I also assume that the firm and the worker can decide to adopt the new technology and invest in training at any point. If training and innovation are adopted at time t^* , productivity is equal to $y + \alpha_1$ in all periods $t > t^*$. Without either training or innovation, output is equal to y .

Let us suppose that all other pairs adopt the new technology at date $t=1$, thus the proportion of firms with the new technology and the proportion of workers with high training are both equal to $\phi_t=1$ for all $t > 1$. Now consider worker i and firm j . First look at the strategy of never adopting the innovation and never training. The return from this strategy is $TS_0 = (1+r)y/r$. If the firm adopts the innovation and trains the worker immediately, the return is $TS_1 = (1+r)y + \alpha_1/r - c - \delta$. If only these two options were being compared, the condition for investment and training to be an equilibrium would be (conditional on coordinating on the better equilibrium): $\alpha_1/r > \delta + c$, the only difference from Section 3.4 being that revenues from innovation and training are accruing for all $t > 1$ here, thus α_1/r rather than $\alpha_1/(1+r)$. I assume in the rest of this section that this inequality holds so that $TS_1 > TS_0$, and I will then show that this is not sufficient to ensure the adoption of the new technology when delay is possible.

When the firm has an option to delay, it can follow the strategy of adopting the new technology whenever it receives a worker who is already trained. The firm will produce y until the worker leaves and a new worker arrives, and this happens with probability s . Since $\phi_t=1$ for all t , the new worker will be trained, thus only the adoption cost has to be incurred.⁷ Therefore, the surplus from that point onwards is $TS_f = (1+r)y + \alpha_1/r - \delta$. The firm will get a proportion $(1-\beta)$ of this amount. Turning to the worker, as soon as he is separated from his initial firm, he will meet a firm with the new technology and the total surplus to be shared at this point between this worker and his new firm is $TS_w = (1+r)y + \alpha_1/r - c$; because the firm already has the innovation, the partnership only needs to pay the cost of training. The worker will receive a proportion β of TS_w . Then, the total surplus of worker i and firm j who choose to delay at time $t=1$ is given as TS_d where

$$TS_d = y + \frac{1}{1+r} [(1-s)TS_d + s(\beta TS_w + (1-\beta)TS_f)]. \quad (4)$$

In words, in every period there is a probability s that the worker and the firm will separate. If separation, occurs, the firm will get $(1-\beta)TS_f$ and the worker will get βTS_w .

7. As it is already assumed that $TS_1 > TS_0$, it is immediate that TS_f is positive, thus the innovation will be adopted. Similarly, TS_w is positive and an untrained worker will be trained immediately by a firm which possesses the new technology.

Substituting for TS_w and TS_f gives

$$TS_d = \frac{(1+r)y}{r} + \frac{s}{r+s} \left(\frac{\alpha_1}{r} - \beta c - (1-\beta)\delta \right). \quad (5)$$

For the innovation to be adopted immediately, it is also necessary that $TS_1 \geq TS_d$. Given $TS_1 > TS_0$, it is immediate that $TS_d > TS_0$. Therefore, the condition for all firms to adopt the new technology and train their workers has become tighter. More specifically, when firms can delay innovation and training decisions, an equilibrium in which all firms invest at $t=1$ requires $\alpha_1 > r(\delta+c) + s(1-\beta)c + s\beta\delta$ which is considerably more restrictive than the condition for innovation and training to be an equilibrium in the absence of the option to delay which was $\alpha_1 > (\delta+c)$.

Intuitively, the skill level of the future workforce of this economy is a *public good* from which all firms benefit (and similarly, the innovation level of firms is a public good from which all workers benefit). When the investment opportunities can be delayed, agents can *free-ride* and reduce their contribution to these public goods by delaying their adoptions and this increases the extent of the inefficiency.⁸

4.2. Opportunistic separations

The analysis in Section 3 assumed that worker i and firm j only separate if they receive the adverse productivity shock. This can be formalized by assuming that both sides incur a cost equal to B when they end their current relation and look for a new partner. If B is sufficiently large, they will not want to separate in the absence of an adverse shock, but when such a shock arrives, they are forced to end the relation. If B is not too large, then the possibility of voluntary separations may affect innovation decisions. In particular, suppose that worker i and firm j who have not invested in the innovation, and all other pairs have. In this case, after a separation worker i will get a firm with the new technology, and firm j will get a more trained worker. Therefore, the incentives not to invest in the first period and separate opportunistically at $t=2$ will be stronger. The previous version of the paper established that with this modification, Proposition 3 still applies as before, but a more stringent condition replaces Condition 1. Condition 2 will remain unchanged since a firm and a worker choosing innovation and high training while others do not will never want to separate voluntarily. Because the new condition is more stringent, it is possible that the multiplicity of equilibria is replaced by a unique equilibrium with low training and no innovation. In any case, inefficiencies are now more serious.⁹

4.3. Unemployment

The analysis so far has not allowed for unemployment. The main reason was to clarify the crucial differences between a competitive economy and one characterized by frictional labour markets. The presence of unemployment does not change any of the key results, but yields some additional insights. To illustrate the effects in the simplest possible way, consider the case in which workers anticipate that if they separate they face a probability

8. Gale (1995) contains a more general analysis of the efficiency consequences of delay in a model with pay-off externalities. Chamley and Gale (1995) illustrate the consequences of delay in a model with informational externalities. Shleifer (1986) shows the possibility of delay in the context of a model with aggregate demand externalities.

9. Details are available from the author upon request.

of unemployment, $p_u > 0$, and that training is useless to an unemployed worker (and also unrealistically that training does not change the probability of unemployment). This formulation immediately implies that a higher probability of unemployment, p_u , which is naturally associated with a higher unemployment rate, reduces the incentives to invest in training, and therefore, the incentives to invest in new technologies.¹⁰

This insight also sheds light on a famous debate started by Habakkuk (1962). Habakkuk claimed that the reason technological progress was faster in the U.S. than in the U.K. during the nineteenth century was the shortage of labour in the U.S. which forced firms to innovate. However, this thesis is not consistent with the fact that most new technologies during this period were not labour saving (see MacLeod (1988)). The model of this paper implies that high unemployment discourages innovation despite the fact that new technologies and labour are complementary, and it offers an explanation for why the lack of labour shortage in nineteenth century Britain may have slowed down labour augmenting technological change.

5. CONCLUSION

This paper analysed innovation and training decisions in a frictional labour market and offered two key results. First, workers do not have the right incentives to invest in general training because they anticipate that part of the productivity gains created by training will be captured by their future employers. This inefficiency in training, in contrast to others proposed in the literature, does not depend on an extreme form of incompleteness of contracts between a worker and his employer, nor does it require credit market imperfections. Also with competitive markets, workers bear the full costs of general training, whereas with labour market frictions, firms may be willing to pay part of the costs of training. Second, if there are also investment decisions in new technology complementary to the skills of the workers, then the inefficiency in training is increased and a multiplicity of equilibria emerges. If a large number of firms are expected to adopt the new technology, workers expect their future productivity to be higher and are more willing to pay for general training. In contrast, there can also be a coordination failure equilibrium where firms do not adopt the innovation and workers choose a low level of skills. The multiplicity of equilibria does not rely on technological spillovers or aggregate demand externalities, it is instead driven by labour market frictions.

APPENDIX: NON-WALRASIAN EQUILIBRIA WITHOUT FRICTIONS

The analysis in subsection 2.3 treated a frictionless labour market as a market coordinated by a Walrasian auctioneer. This is not the only possible interpretation. The crucial imperfection in the economy of Section 3 is that workers and firms cannot costlessly change partners (within a period). Thus a natural benchmark is the limit point where there are no costs of changing partners. However, this is not necessarily the same as a Walrasian equilibrium. Instead it is the (subgame perfect) equilibrium of a game in which firms compete *a la* Bertrand in each period for all the workers. Let me refer to this equilibrium as a *weak competitive equilibrium*¹¹ and characterize these equilibria in this Appendix.

10. Also this channel can lead to a further multiplicity of equilibria. In one equilibrium, unemployment is high, therefore workers do not want to invest in skills, this reduces profitability of firms, and as a result, few firms enter, supporting an equilibrium with high unemployment. Details, which were included in previous versions, are available upon request.

11. This is an equilibrium concept first proposed by Hart (1979) and Makowski (1980) in the context of monopolistic competition, and referred to by Allen and Gale (1988) as *conjectural equilibrium*. I use the term *weak competitive equilibrium* to emphasize the fact that this concept is used implicitly as the competitive equilibrium concept but it is weaker than the *Walrasian equilibrium*.

The game underlying the *weak competitive equilibrium* is as follows: at all times $t \geq 2$, firms hire labour with training τ at the wage $w_t(\tau): [0, \bar{\tau}] \rightarrow \mathbb{R}^+$. Then all agents take the sequence of future equilibrium wage functions $\{w_t(\tau)\}$ as given, and workers are hired at time $t = 1$ at wage $v(\tau): [0, \bar{\tau}] \rightarrow \mathbb{R}$ where τ is the amount of training that the employer will provide during this period. An equilibrium is a sequence of functions $\{w_t(\tau)\}_{t=2}^T$ and $v(\tau)$ such that all agents' demands are satisfied at the time t spot labour market for all t .¹² Note the difference between this equilibrium concept and the Walrasian one: here the market for worker with training τ only needs to clear if there are workers of training level τ . In contrast, with the Walrasian equilibrium, even if there are no workers with training level τ , we still have to find a price $w(\tau)$ such that at this price, no firm wants to hire workers with training level τ and no worker wants to obtain training τ rather than some other level. Thus, it must be clear that a *Walrasian equilibrium* is a *weak competitive equilibrium* but not vice versa.

Let me start with the last period of this economy, $t = T = 2$, and denote the proportion of firms with $\gamma = 1$ (i.e. the firms with the new technology) by ϕ . There will be some pairs of workers and firms who cannot work together [that is, if worker i and firm j were together in period $t = 1$ and received the adverse shock], but this will not create any problem unless $\phi = 0$ or $\phi = 1$ because there will be a continuum of firms that can employ each worker. Before the next result, I also define, as in the text, $q(\bar{\tau})$ as the distribution function of workers with training $\bar{\tau}$ in period $t = 2$.

Lemma 1. *Suppose $0 < \phi < 1$, then in all $t = 2$ weak competitive equilibria, workers with $r \geq \tau^*$ work with firms that have $\gamma = 1$ and workers with $\tau < \tau^*$ work with firms that have $\gamma = 0$ where τ^* is such that $q(\tau^*) = 1 - \phi$. For any training level τ at which there exist a positive measure of workers, wages are given by*

$$\begin{aligned} w(\tau) &= y + \alpha_1 \tau - a && \text{if } \tau \geq \tau^* \\ w(\tau) &= y + \alpha_0 \tau - b && \text{if } \tau < \tau^*, \end{aligned} \tag{6}$$

where $a - (\alpha_1 - \alpha_0)\tau^* = b$.

Proof. The proof is by contradiction. I first prove that no other allocation rule can be equilibrium, then given the allocation rule, I show that only the wage function in (6) is equilibrium.

Suppose that a worker with $\tau' > \tau^*$ is employed by a firm without the innovation, j' . Let the profit of this firm be $\pi(j')$, then $\pi(j') = y + \alpha_0 \tau' - w(\tau')$. Let the profit of this firm from hiring worker τ^* be $\pi^*(j')$, then $\pi^*(j') = y + \alpha_0 \tau^* - w(\tau^*)$. For equilibrium, it is necessary that $\pi(j') = y + \alpha_0 \tau' - w(\tau') \geq \pi^*(j') = y + \alpha_0 \tau^* - w(\tau^*)$ which implies that $w(\tau') - w(\tau^*) \leq \alpha_0(\tau' - \tau^*)$. Suppose this last inequality is true. Then because a firm with the innovation must be indifferent between hiring τ^* and τ' , $w(\tau') - w(\tau^*) \geq \alpha_1(\tau' - \tau^*)$ which, since $\alpha_1 > \alpha_0$, gives a contradiction. This implies that all workers with $\tau \geq \tau^*$ are hired by firms that have $\gamma = 1$, and therefore, all workers with $\tau < \tau^*$ must be hired by firms with $\gamma = 0$.

Now take this allocation rule as given. Take two training levels $\tau'' > \tau' \geq \tau^*$. Then firms with the innovation must be indifferent between hiring any of these two workers, thus $w(\tau'') - w(\tau') = \alpha_1(\tau'' - \tau')$ which gives the first part of (6). Next consider two workers with $\tau' < \tau'' < \tau^*$. The same reasoning implies $w(\tau'') - w(\tau') = \alpha_0(\tau'' - \tau')$. Thus the second part of (6) is obtained. Now consider $\tau'' \geq \tau^* > \tau'$. Since a firm with $\gamma = 1$ is hiring τ'' , we require $\pi_1(\tau'') \geq \pi_1(\tau')$ where $\pi_1(\tau)$ denotes the profit that a firm with the new technology makes from a worker with training equal to τ . Thus, it is necessary that $w(\tau'') - w(\tau') \leq \alpha_1(\tau'' - \tau')$. This implies that $b \leq a - (\alpha_1 - \alpha_0)\tau'$. Since this has to be true for all $\tau' \leq \tau^*$, then we also have $b \leq a - (\alpha_1 - \alpha_0)\tau^*$. Next a firm without the technology can always hire τ^* , and obtain profits equal to $\pi_0(\tau^*) = a_0 \tau^* - w(\tau^*) = a - (\alpha_1 - \alpha_0)\tau^*$. This firm is currently making profits equal to $\pi_0(\tau') = b$, therefore, $\pi_0(\tau') \geq \pi_0(\tau^*)$, hence $b \geq a - (\alpha_1 - \alpha_0)\tau^*$. Combining this with the previous inequality, we get $b = a - (\alpha_1 - \alpha_0)\tau^*$, which completes the proof. \parallel

An important result contained in this lemma is that the allocation of workers to firms will be optimal; high-training workers will be employed by firms that have the new technology. Moreover, irrespective of the skill distribution of workers, firms with the new technology will pay the marginal product of a worker, thus they will be indifferent to the exact training of the worker they hire in equilibrium. Also note the statement in the lemma which requires some positive measure of workers to have training τ for the above wage rules to apply. This requirement is not necessary when we are considering a Walrasian equilibrium. But as noted above, with the weaker definition, only commodities which are actively traded can be priced.

12. Since firms are competing *a la* Bertrand, the requirement that markets clear, and the offers are best-response to each other are equivalent.

Also, the above lemma is for $0 < \phi < 1$. If instead $\phi = 0$, there is no demand from firms that have the new technology, thus wages are simply given as

$$w(\tau) = y + \alpha_0 \tau - \tilde{a},$$

and similarly when $\phi = 1$, $w(\tau) = y + \alpha_1 \tau - \hat{a}$, for \tilde{a} and \hat{a} non-negative.

Proposition 4. *Assume that Assumption 3 holds, then there exist exactly two weak competitive equilibrium allocations.*

- (i) $\phi = 1$, and all workers choose τ^h .
- (ii) $\phi = 0$, and all workers choose τ^l .

Proof. Part A shows that (i) above is an equilibrium. Part B shows that (ii) is an equilibrium and Part C shows that there are no others.

Part A: Suppose all firms choose to invest in the new technology so that $\phi = 1$. Then $w(\tau) = y + \alpha_1 \tau - a$, and each worker knowing that his wages will be higher by $\alpha_1 \tau$ is willing to pay up to the point where $\alpha_1 = (1+r)c'(\tau)$ which gives the training level equal to τ^h (and workers pay $c(\tau^h)$ in the first period to the firm to compensate it for the cost of training). Since all workers have training τ^h , the profit levels to the alternative strategies of a firm are given as $\pi_1 = a - (1+r)\delta$ and $\pi_0 = a - (\alpha_1 - \alpha_0)\tau^h$. Assumption 3(i) implies that $\alpha_1 \tau^h - \alpha_0 \tau^l > (1+r)[c(\tau^h) - c(\tau^l) + \delta]$. Given the strict convexity of $c(\cdot)$, it is also the case that $(1+r)[c(\tau^h) - c(\tau^l)] > \alpha_0(\tau^h - \tau^l)$, thus $\alpha_1(\tau^h - \tau^l) > (1+r)\delta$. Hence (i) above is an equilibrium.

Part B: Now suppose all firms choose $\gamma = 0$, then second period wages are $w(\tau) = y + \alpha_0 \tau - \tilde{a}$ because there is no firm with the advanced technology that may get additional marginal product α_1 from training. Anticipating this wage function, workers choose τ^l . With a similar argument to Part A above, it follows from Assumption 3(ii) that firms are happy to choose $\gamma = 0$.

Part C: Next suppose that there is an equilibrium with $\phi \in (0, 1)$. Then the wage function in Lemma 1 applies. Therefore, $\pi_1 = a$ and $\pi_0 = b$. If $a - b < (1+r)\delta$, then the firms that choose $\gamma = 1$ can increase their profits by not investing. If $a - b > (1+r)\delta$, then firms with $\gamma = 0$ can increase their profits by investing. Finally $a - b = (1+r)\delta$ can only be an equilibrium if a proportion $(1 - \phi)$ of the workers choose τ^l . In this equilibrium, $w(\tau^h) - w(\tau^l) = \alpha_1(\tau^h - \tau^l) - (1+r)\delta$, thus the net return to choosing τ^h rather than τ^l is $\alpha_1(\tau^h - \tau^l) - (1+r)\delta - (1+r)[c(\tau^h) - c(\tau^l)]$, which by Assumption 3(i) is positive, thus no worker chooses τ^l , hence there are no equilibria with $\phi \in (0, 1)$. ||

This proposition establishes that as well as the Pareto optimal equilibrium there is a coordination-failure equilibrium. In this second equilibrium, (ii), investment in the new technology is not undertaken and a lower than socially optimal level of training is chosen. The existence of this ‘‘coordination-failure’’ equilibrium is due to the fact that the right markets are not open at time $t = 2$. In particular, when all workers choose τ^l , there is no one hiring or supplying labour in the market for workers with training τ^h , thus this market is closed and the appropriate price signals are not being transmitted. In contrast, suppose the wage determination rule in the lemma applies to τ^h even when there are no workers with training τ^h , then it is clear that workers would prefer to obtain training τ^h rather than τ^l . In other words, in a Walrasian equilibrium, if the market for workers with training τ^h is closed, then their price (wage) must be zero and there must be excess supply at this price. Yet at zero price, firms will naturally want to hire workers with training level τ^h at $t = 2$ [see also the discussion in Allen and Gale (1988)]. Therefore, if labour with training τ^h were priced (even when there are no workers at this training level), the allocation with τ^l and $\phi = 0$ would not be an equilibrium. This reasoning immediately establishes Proposition 1 in the text: there is a unique Pareto optimal Walrasian equilibrium.

Nevertheless, the concept of *weak competitive equilibrium* may appear more relevant than the abstract notion of *Walrasian equilibrium* which seems to demand even more than usual from the auctioneer (i.e. to price commodities not traded along the ‘‘equilibrium path’’). In fact, in many economic applications, the concept of weak competitive equilibrium is used implicitly rather than the Walrasian equilibrium. However, the allocation with $\gamma = 0$ and $\tau = \tau^l$ is not even a robust weak competitive equilibrium. In a game-theoretic sense it is ‘‘unstable’’ because it does not survive arbitrarily small perturbations.

To demonstrate this in the simplest way, consider an economy E_ε which is the same as the economy above except that firms make a mistake/tremble with probability ε and choose the opposite of the investment choice they intended (e.g. they invest if they intended not to invest). Then:

Definition 1. Let A be an equilibrium allocation and $\tau_A: [0, 1] \rightarrow \mathbb{R}$ be a mapping that allocates a training level to each worker in this equilibrium. A satisfies *stability* iff $\forall v > 0, \exists \bar{\varepsilon} > 0$ such that $\forall \varepsilon < \bar{\varepsilon}, A_\varepsilon$ is an equilibrium of E_ε where in A_ε training levels are given by τ_ε such that $\|\tau_\varepsilon - \tau_A\| < v$.

In words, an equilibrium satisfies stability if a small change in the behaviour of the firms leads to a small change in the behaviour of the workers.¹³

Proposition 5. *Suppose Assumption 3 holds, then the allocation $\phi=1$ and $\tau=\tau^h$ is the unique weak competitive equilibrium which satisfies stability.*

Proof. It is necessary to establish that for all $\nu>0$, there exists τ_ν that is best response to $\phi=1-\varepsilon$ for all $\varepsilon<\bar{\varepsilon}$ such that $\|\tau^h-\tau_\nu\|\leq\nu$. Let $w(\tau^h)=y+\alpha_1\tau^h-a$ as in Lemma 1. A worker who deviates and chooses $\tau<\tau^h$ will have the lowest training level and thus by Lemma 1 will be matched with the firms that have $\gamma=0$. He will therefore have a wage rate $w(\tau)=y+\alpha_1\tau-b$ for all $\tau<\tau^h$. Since the marginal return to τ is still α_1 , $\tau=\tau^h$ is the best-response, thus $\tau_\nu=\tau^h$. Therefore, the requirement is satisfied for all $\nu>0$, and the equilibrium allocation with $\phi=1$ and $\tau=\tau^h$ satisfies stability.

Next consider the equilibrium allocation with $\phi=0$ and $\tau=\tau'$. From Lemma 1, $\forall\phi=\varepsilon>0$, by the same argument as in Part C of Proposition 4, $\tau=\tau^h$ is best-response, thus $\tau_\nu=\tau^h$. So let $\nu<\tau^h-\tau'$, then $\|\tau'-\tau_\nu\|>\nu$ and this allocation does not satisfy stability. \parallel

The intuition of this result is quite simple. The coordination failure equilibrium was pathological as it relied on one of the markets not transmitting price signals. As soon as there is some activity in all markets such that even commodities not traded along the equilibrium path have the right prices, this equilibrium disappears.

Therefore, the analysis of this Appendix confirms the conclusion reached in subsection 2.3 that in the absence of labour market frictions, there is a unique efficient equilibrium.

It is also straightforward to see that the ‘‘coordination failure’’ equilibrium of Section 3 when labour markets are imperfect satisfies the notion of stability introduced in this Appendix. A small fraction of firms investing in the new technology would only have a minimal impact on the expected return to training, and thus workers will only respond by a very small amount to such changes.

Acknowledgements. I am grateful to Charlie Bean, Peter Diamond, Oliver Hart, Steve Pischke, Kevin Roberts, Jim Robinson, Andrew Scott and three anonymous referees. I also thank seminar participants at Boston College, Chicago, LSE, Northwestern, MIT, Princeton, UCL and Yale.

REFERENCES

- ACEMOGLU, D. (1996a), ‘‘A Microfoundation For Increasing Returns in Human Capital Accumulation’’, *Quarterly Journal of Economics*, **111**, 779–804.
- ACEMOGLU, D. (1996b), ‘‘Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence’’ (CEPR Discussion Paper, No. 1459).
- ACEMOGLU, D. (1997), ‘‘Matching, Heterogeneity and the Evolution of Income Distribution’’, *Journal of Economic Growth*, **1**, 40–65.
- ACEMOGLU, D. and PISCHKE, S. (1996), ‘‘Why Do Firms Train? Theory and Evidence’’ (National Bureau of Economic Research, Working Paper No. 5605).
- ALLEN, F. and GALE, D. (1988), ‘‘Optimal Security Design’’, *Review of Financial Studies*, **1**, 229–263.
- ANDERSEN, P. M. and MEYER, B. D. (1994), ‘‘The Extent and Consequences of Job Turnover’’, *Brookings Papers on Economic Activity, Microeconomics*, 177–248.
- BARTEL, A. and LICHTENBERG, F. (1987), ‘‘The Comparative Advantage of Educated Workers in Implementing New Technologies’’, *Review of Economics and Statistics*, **69**, 1–11.
- BECKER, G. (1964) *Human Capital* (Chicago: University of Chicago Press).
- BINMORE, K., RUBINSTEIN, A. and WOLINSKY, A. (1986), ‘‘The Nash Bargaining Solution in Economic Modelling’’, *Rand Journal of Economics*, **17**, 176–188.
- BISHOP, J. (1991), ‘‘On-the-Job Training of New Hires’’, in Stern, D. and Ritzen, J. M. (eds.), *Market Failure in Training?* (Berlin: Springer-Verlag).
- BLANCHARD, O. and DIAMOND, P. (1990), ‘‘The Cyclical Behavior of the Gross Flows of U.S. Workers’’, *Brookings Papers On Economic Activity*, **2**, 85–155.

13. An alternative definition of stability can be given in terms of the slopes of best-response functions, see for instance Fudenberg and Tirole (1991) pp. 23–28). The notion of stability used here is related to refinements such as Trembling Hand Perfection and is therefore weaker than the definition of asymptotic stability given there, and thus, the result that the coordination failure equilibrium is not stable derived here is a stronger result. In particular, it is straightforward to see that the best-response functions of workers in this case are discontinuous: when no firm adopts the innovation, they all want to receive training equal to τ' but as soon as there is even a very small proportion of firms adopting the innovation, all workers would like to obtain τ^h . Thus, as soon as there is a very small number of firms who demand the new technology, the best-response of the workers jumps from τ' to τ^h .

- BLINDER, A. S. and KRUEGER, A. B. (1991), "International Differences in Turnover: A Comparative Study with Emphasis on the U.S. and Japan" (mimeo, Princeton University).
- BURDETT, K. and COLES, M. (1995), "Marriage and Class" (mimeo, University of Essex).
- BURDETT, K. and SMITH, E. (1996), "Education and Matching Externalities", in Booth, A. and Snower, D. (eds.), *Skills Gap and Economic Activity* (Cambridge: Cambridge University Press).
- CHAMLEY, C. and GALE, D. (1994), "Information Revelation and Strategic Delay in a Model of Investment", *Econometrica*, **62**, 1065–1085.
- COOPER, R. and JOHN, A. (1988), "Coordinating Coordination Failures in Keynesian Models", *Quarterly Journal of Economics*, **103**, 441–465.
- DIAMOND, P. (1982), "Aggregate Demand Management in a Search Equilibrium", *Journal of Political Economy*, **90**, 881–894.
- DIAMOND, P. and MASKIN, E. (1979), "An Equilibrium Analysis of Search and Breach of Contract, I: Steady States", *Bell Journal of Economics*, **10**, 282–316.
- DURLAUF, S. (1993), "Non-ergodic Growth", *Review of Economic Studies*, **60**, 349–366.
- FUDENBERG, D. and TIROLE, J. (1991) *Game Theory* (Cambridge: MIT Press).
- GALE, D. (1995), "Dynamic Coordination Games", *Economic Theory*, **5**, 1–18.
- GRILLICHES, Z. (1969), "Capital and Skill Complementarity", *Review of Economics and Statistics*, **51**, 465–468.
- GROUT, P. (1984), "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach", *Econometrica*, **52**, 449–460.
- HABAKKUK, H. J. (1962) *American and British Technology in the Nineteenth Century* (Cambridge: Cambridge University Press).
- HARHOFF, D. and KANE, T. J. (1994), "Financing Apprenticeship Training: Evidence from Germany" (NBER Working Paper No. 4557).
- HART, O. D. (1979), "Monopolistic Competition in a Large Economy With Differentiated Commodities", *Review of Economic Studies*, **46**, 1–30.
- HASHIMOTO, M. (1991), "Training and Employment Relations in Japanese Firms", in Stern, D. and Ritzen, J. M. (eds.), *Market Failure in Training?* (Berlin: Springer-Verlag).
- JONES, I. (1985), "Apprentice Training Costs in British Manufacturing Establishments: Some New Evidence", *British Journal of Industrial Relations*, **24**, 333–362.
- KIYOTAKI, N. (1988), "Multiple Expectational Equilibria under Monopolistic Competition", *Quarterly Journal of Economics*, **103**, 695–713.
- LAING, D., PALIVOS, T. and WANG, P. (1995), "Learning, Matching and Growth", *Review of Economic Studies*, **62**, 115–131.
- LEONARD, J. S. (1987), "In the Wrong Place at the Wrong Time: The Extent of Frictional and Structural Unemployment", in Lang, K. and Leonard, J. S. (eds.), *Unemployment and the Structure of the Labor Markets* (Oxford: Basil Blackwell).
- MACLEOD, C. (1988) *Inventing the Industrial Revolution* (Cambridge: Cambridge University Press).
- MAKOWSKI, L. (1980), "Perfect Competition, the Profit Criterion and the Organization of Economic Activity", *Journal of Economic Activity*, **22**, 222–242.
- MORTENSEN, D. (1982), "Property Rights and Efficiency in Mating, Racing and Related Games", *American Economic Review*, **72**, 968–979.
- MURPHY, K., SHLEIFER, A. and VISHNY, R. (1989), "Industrialization and the Big Push", *Journal of Political Economy*, **97**, 1003–1026.
- NORTHCOTT, J. and WALLING, A. (1988), "The Impact of Microelectronics; Diffusion Benefits and Problems in British Industry (Policy Studies Institute, London).
- NORTHCOTT, J. and VICKREY, G. (1993), "Surveys of the Diffusion of Microelectronics and Advanced Manufacturing Technology" (Paper presented at MIT/NSF/OECD Workshop on "The Productivity Impact of Information Technology Investments").
- OECD (1994) *Employment Outlook*.
- PISCHKE, S. (1996), "Continuous Training in Germany" (NBER Working Paper No. 8529).
- PISSARIDES, C. (1990) *Equilibrium Unemployment Theory* (Oxford: Basil Blackwell).
- ROBINSON, J. A. (1995), "Unemployment and Human Capital Accumulation" (mimeo, USC).
- ROSENSTEIN-RODAN, P. (1943), "Problems of Industrialization of Eastern and Southern-Eastern Europe", *Economic Journal*, **53**, 202–211.
- SAINT-PAUL, G. (1996), "Unemployment, Wage Rigidity and the Returns to Education", *Journal of Public Economics* (forthcoming).
- SHAKED, A. and SUTTON, J. (1984), "Involuntary Unemployment as a Perfect Equilibrium in a Bargaining Model", *Econometrica*, **52**, 1341–1364.
- SHLEIFER, A. (1986), "Implementation Cycles", *Journal of Political Economy*, **94**, 1163–1190.
- TAN, H. W. (1991), "Technical Change and Human Capital Acquisition in the U.S. and Japanese Labor Markets", in Hulten, C. R. (ed.), *Productivity Growth in Japan and the United States* (Chicago: University of Chicago Press).