OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 68, 4 (2006) 0305-9049

# The Impact of Training on Productivity and Wages: Evidence from British Panel Data*

LORRAINE DEARDEN,† HOWARD REED‡ and JOHN VAN REENEN§

*†Institute for Fiscal Studies and Institute of Education, London, UK (e-mail: ldearden@ifs.org.uk)*
*‡Institute for Public Policy Research, London, UK (e-mail: h.reed@ippr.org)*
*§Centre for Economic Performance, London School of Economics and CEPR, London UK (e-mail: j.vanreenen@lse.ac.uk)*

## Abstract

It is standard in the literature on training to use wages as a sufficient statistic for productivity. This paper examines the effects of work-related training on direct measures of productivity. Using a new panel of British industries 1983–96 and a variety of estimation techniques we find that work-related training is associated with significantly higher productivity. A 1% point increase in training is associated with an increase in value added per hour of about 0.6% and an increase in hourly wages of about 0.3%. We also show evidence using individual-level data sets that is suggestive of training externalities.

*Financial support from the Economic and Social Research Council and the Leverhulme Trust is gratefully acknowledged. We would like to thank Fernando Galindo-Ruedo, Amanda Gosling, Richard Layard, Lisa Lynch, Stephen Machin, Steve McIntosh, Alan Manning, Steve Pischke, two anonymous referees and participants in many seminars for helpful comments. We are grateful to the Office for National Statistics (ONS) for supplying the Census of Production data and to the Organization for Economic Co-operation and Development (OECD) for supplying the Inter-Sectoral Data Base (ISDB) data. We would like to thank Martin Conyon for supplying the firm-level training data and Jonathan Haskel for providing us with some of the industry price indices used in this paper. Material from the Labour Force Survey is Crown Copyright, has been made available by the ONS through the ESRC Data Archive and has been used by permission. Neither the ONS nor the Data Archive bears any responsibility for the analysis or interpretation of the data reported here. An earlier version of this paper (with a longer data description) has appeared as 'Who gains when workers train', Working Paper No. 00/04, Institute for Fiscal Studies.

JEL Classification numbers: C23, D24, J31.

## I. Introduction

It is a widely held view that Britain needs to increase work-related training to improve long-term economic performance and address the 'skills gap'.[1] Despite the policy interest and the huge economics literature on human capital, there are hardly any papers that examine the impact of work-related training on direct measures of productivity. The primary contribution of our paper is to provide such evidence for the first time in the UK and for the first time anywhere over a long period (we have 14 years of data). Analysis of the impact of training on productivity has focused almost entirely on estimating the impact of training on wages. Most studies looking at the private return to work-related training find that training results in workers receiving higher real wages.[2]

Although these studies are informative, they only tell half the story as they ignore the impact on the employer's productivity. The relationship between wage increases and productivity gains can vary according to the structure of the labour and product markets and according to who actually pays the costs of training. In the simplest neoclassical view of the labour market where the market is perfectly competitive, wages will be equal to the value of marginal product. Thus the wage can be taken as a direct measure of productivity. This simple relationship can break down for many reasons. For example, in Becker's model of specific human capital, the employer will pay for training, so there should be no effect of completed training spells on observed wages even though there may be a large impact on productivity.[3]

If the labour market is characterized by imperfect competition then the strict link between wages and productivity is usually broken. Employees can find themselves being paid less (or more) than their marginal revenue product. Nevertheless, it is still the case that conditional on a given degree of rent-sharing or monopsony power; increases in wages have to be paid out of productivity gains. Therefore, we can assert the general principle that these

[1]See Green and Steedman (1997) or National Skills Task Force (1998). In the December 2003 Pre-Budget Report, the British Chancellor justified the extension of the Employer Training Pilots in order to help improve the skills gap and UK productivity (http://www.hm-treasury.gov.uk/media//2E3BD/03_Meeting%20the%20Pro_EF.pdf).

[2]See Greenhalgh and Stewart (1987), Booth (1991, 1993) or Blundell, Dearden and Meghir (1996) for UK evidence. US studies using panel data include Lillard and Tan (1992), Lynch (1992), Blanchflower and Lynch (1992) and Bartel and Sicherman (1999). Winkelmann (1994) uses German data and Bartel (1995) looks within a large US manufacturing company.

[3]There are many other reasons for a wedge between productivity and wage in a competitive labour market. First, employees may receive non-pecuniary benefits from training. Secondly, workers may implicitly pay the costs of a training scheme in the form of lower wages whilst being trained, which then rise after training is completed – so we might see a greater increase in *observed* wages than in productivity. Thirdly, employees' wages could be lower during training because they are not contributing to firm productivity whilst actually being trained. Fourthly, there may be deferred compensation packages where the employee's remuneration is 'backloaded' towards later post-training years as a means of ensuring loyalty and/or effort early in the employee's tenure (e.g. Lazear, 1979).

real wage increases should provide a lower bound on the probable size of productivity increases. In practice, the productivity gains are likely to be higher than this. For instance, in a labour market with frictions and some wage compression (e.g. from a binding minimum wage), there will be productivity gains even from general training that are not passed on to the employee in terms of wages but are only reflected in direct measures of productivity.[4] Similar results can be found in some bargaining models (e.g. Booth, Francesconi and Zoega, 1999).

There exist a small number of empirical papers that relate firm productivity to a measure of training.[5] Although a positive correlation is generally found, it is very difficult to interpret because the training measures are only measured at a single point of time and could be picking up many unobservable firm-specific factors correlated with both training and productivity. Black and Lynch (2001) used an establishment training survey at two points of time. In the cross-section, they identified some effects of the type of training on productivity, but they found no significant association when they controlled for plant-specific effects. Ichniowski, Shaw and Prennushi (1997) investigated the factors that influence productivity in a panel of US steel finishing mills. After controlling for fixed effects, they found a role for training only in combination with a large variety of complementary human resource practices. Carriou and Jeger (1997), Ballot, Fakhfakh and Taymaz (1998) and Delame and Kramarz (1997) used French firm-level panel data to look at the effects of training on value added and found positive and significant effects. Although these studies are broadly consistent with our own, they do not fully exploit the potential of their panel data by allowing training to be a choice variable.[6]

Our contribution in this paper is to advance the literature in at least three ways. Black and Lynch (2001) emphasize the problems of trying to identify the effects of training in a short panel (they have only two separate training observations). Although unobserved heterogeneity can be controlled for through fixed effects with only two periods, attenuation biases due to measurement error are exacerbated. To address this, we build a panel that contains up to 14 consecutive years of training data. Secondly, we explicitly allow training to be a choice variable by using general method of moments (GMM) estimators developed to deal with endogenous variables in production functions. Thirdly, we combine estimation of the productivity effects of

[4]See Acemoglu and Pischke (1999, 2003).

[5]Black and Lynch (1996), Bartel (1994), Barrett and O'Connell (2001), De Koning (1994), Boon and van der Eijken (1997) and Ballot and Taymaz (1998) have objective productivity measures. Bartel (1995), Holzer (1990), Barron, Black and Zoewenstein (1989) and Krueger and Rouse (1998) use subjective measures of productivity. Holzer *et al.* (1993) do find effects of changes in productivity on changes in one measure of quality – the scrap rate.

[6]See Greenhalgh (2002) for a much more extensive review of the French and UK literature in this area.

training with estimation of the wage effects of training. Although comparisons between the production function and the wage equation are becoming more common for other worker characteristics such as gender and human capital, this is the first time the strategy has been used for training.[7] In principle, this allows us to examine whether trained workers are paid the value of their marginal product.

We conduct our main analysis of the effects of training at the industry level (although we have also performed estimation at the firm and individual level for comparative purposes). There is simply no alternative to this strategy if one wishes to use long time series of training and productivity information. The only publicly available firm-level panel data in the UK is a sample of about 119 firms in the late 1990s with only very basic training information (a concise investigation of this is presented in Dearden, Reed and Van Reenen, 2005, Appendix B). Aggregation has pros and cons that are discussed in this paper. On the positive side, if there are important spillovers to training within an industry (e.g. through a faster rate of innovation) then a firm-level analysis will potentially miss out these linkages and underestimate the return to human capital.[8] On the negative side, there may be aggregation biases at the sectoral level that could lead to negative or positive biases on the training coefficient. We follow Grunfeld and Griliches (1960) in arguing that the pros of aggregation probably outweigh the cons.

The format of the paper is as follows. Section II describes the simple economic models of productivity and wages that we will estimate and section III details the econometric strategy. The data are described in section IV and the results are presented in section V. Section VI offers some concluding comments. Dearden *et al.* (2005) contains more information on the data and some additional experiments. Our main result is that we find a statistically and economically significant effect of training on industrial productivity. A 1% point increase in training is associated with about a 0.6% increase in productivity and a 0.3% increase in hourly wages. The productivity effect of training is twice as large as the wage effect, implying that existing studies have underestimated the benefits of training by focusing on wages.

[7]Hellerstein, Neumark and Troske (1999), Hellerstein and Neumark (1999), Hægeland and Klette (1999) and Jones (2001) examine the differential impact of human capital and gender on wages and productivity. A recent study that utilizes our methodology and looks at this question using a panel of French and Swedish firms is Ballot, Fakhfakh and Taymaz (2002). They find that both French and Swedish firms appropriate a high proportion of the returns to training (82% and 67% respectively).

[8]For example, O'Mahony (1998) finds that the coefficient on labour skills in a production function is more than twice that assumed by traditional growth accounting from relative wages. Other recent papers that have looked at the impact of human capital on directly measured productivity include Moretti (2004) on US data and Haskel, Hawkes and Pereira (2003) on UK data.

## II. A model of training and productivity

To see our approach, assume that we can characterize a representative plant in an industry by a Cobb–Douglas production function written in value added form[9]

$$Q = AL^{\alpha}K^{\beta}, \tag{1}$$

where $Q$ is value added, $L$ is effective labour input allowing for quality and quantity dimensions, $K$ is capital and $A$ is a Hicks neutral efficiency parameter.

We consider that trained workers are more productive than untrained workers, so that effective labour input can be written as

$$L = N^{\mathrm{U}} + \gamma N^{\mathrm{T}}, \tag{2}$$

where $N^{\mathrm{U}}$ is the number of untrained workers, $N^{\mathrm{T}}$ is the number of trained workers and $\gamma$ is a parameter which, if trained workers are more productive than non-trained workers, will be greater than unity. The total number of workers, $N$, is equal to the sum of trained and untrained workers. Substituting equation (2) into equation (1) gives

$$Q = A[1 + (\gamma - 1)\mathrm{TRAIN}]^{\alpha}N^{\alpha}K^{\beta}, \tag{3}$$

where $\mathrm{TRAIN} = N^{\mathrm{T}}/N$ is the proportion of trained workers in an industry. Taking natural logarithms, we obtain

$$\ln Q = \ln A + \alpha \ln[1 + (\gamma - 1)\mathrm{TRAIN}] + \alpha \ln N + \beta \ln K. \tag{4}$$

This could be estimated by non-linear least squares. If $(\gamma - 1)$ TRAIN is 'small', we can use the approximation $\ln(1 + x) = x$ and rewrite the production function as[10]

$$\ln Q = \ln A + \alpha(\gamma - 1)\mathrm{TRAIN} + \alpha \ln N + \beta \ln K. \tag{5}$$

If the industry exhibits constant returns to scale (i.e. $\alpha + \beta = 1$) then equation (5) can be rewritten in terms of labour productivity as

$$\ln\left(\frac{Q}{N}\right) = \ln A + (1 - \beta)(\gamma - 1)\mathrm{TRAIN} + \beta \ln\left(\frac{K}{N}\right). \tag{6}$$

If the trained are no more productive than the untrained ($\gamma = 1$) then the coefficient on TRAIN will be zero.

This method can be easily extended to a larger number of different types of heterogeneous workers in the labour quality index. If we index the discrete

[9]This should be viewed as a first-order approximation to a more complicated functional form. It is straightforward to generalize this to more complex functional forms such as translog and some experiments are included in the empirical results.

[10]The results were estimated both by non-linear least squares and by least squares using the approximation. The results did not significantly differ (see Table 3), so the more convenient log-linear approximation is used for the baseline results.

type of labour by $k$ (where until now we have discussed $k$ solely in terms of the training status of workers) then equation (4) can be written

$$\ln Q = \ln A + \alpha \ln\left\{1 + \sum_k \left[(\gamma_k - 1)\left(\frac{N_k}{N}\right)\right]\right\} + \alpha \ln N + \beta \ln K. \qquad (7)$$

Empirically, we will allow for many other dimensions of labour quality such as education, occupation, age, tenure and gender.[11] There are a large number of other influences on productivity captured in $A$, so we allow for differential hours, worker turnover rates, innovation (as proxied by research and development expenditures), regional composition and the proportion of small firms. Labelling these factors as $X$, imposing constant returns and using the log approximation, the basic production function becomes

$$\ln\left(\frac{Q}{N}\right) = (1-\beta)\sum_k \left[(\gamma_k - 1)\left(\frac{N_k}{N}\right)\right] + \beta \ln\left(\frac{K}{N}\right) + \delta' X. \qquad (8)$$

The wage equation that we estimate parallels the productivity equation in (8). We view the wage equation as more of a descriptive regression than the structurally derived production function. Under competitive spot markets for labour, relative wages should equal the relative marginal productivities of workers of different types. This is because if the relative productivity of trained workers, $\gamma$, exceeded the relative wages of trained workers then employers would only employ trained workers (Hellerstein *et al.*, 1999).

Consider the wage bill, $W$, for the representative plant in an industry. Again, take the simplest model where there are only two types of workers: trained workers paid average wage $w^{\mathrm{T}}$ and untrained workers paid average wage $w^{\mathrm{NT}}$. Relative wages are $\lambda = w^{\mathrm{T}}/w^{\mathrm{NT}}$. By definition,

$$W = w^{\mathrm{NT}}(N - N^{\mathrm{T}}) + \lambda w^{\mathrm{NT}} N^{\mathrm{T}} = w^{\mathrm{NT}}[N + (\lambda - 1)N^{\mathrm{T}}]. \qquad (9)$$

In logarithms, the average wage ($w$) is

$$\ln w = \ln\left(\frac{W}{N}\right) = a + \ln[1 + (\lambda - 1)\mathrm{TRAIN}], \qquad (10)$$

where $a = \ln(w^{\mathrm{NT}})$.

Clearly, estimation of equation (10) can be used to recover the relative wage mark-up associated with training, $\lambda$, and then compared with the relative productivity effect of training, $\gamma$. Parallel to the productivity equation, we will

[11] We follow Hellerstein *et al.* (1999) by entering these variables in linear proportions. One could allow a larger number of cells for interactions of the labour quality variables (e.g. the proportion of educated men – a two-way interaction – or the proportion of educated men who are trained – a three-way interaction). We experimented with some breakdowns like this on the training variable, but Labour Force Survey (LFS) cell sizes by industry were generally not large enough.

allow for multiple types of labour quality, capital and other factors to influence wages. The empirical wage equation to be estimated is therefore:[12]

$$\ln w = a + \sum_k \left[ (\lambda_k - 1)\left(\frac{N_k}{N}\right) \right] + \beta^w \ln K + \delta^{w\prime} X. \qquad (11)$$

## III. Econometric modelling strategy

The basic equation we estimate can be written in simplified form as

$$y_{it} = \theta x_{it} + u_{it}, \qquad (12)$$

where $y$ is $Q/N$ and $x$ is a vector of (suspected endogenous) variables including training. Subscript $i$ indicates the representative firm in an industry, $t$ is time and $\theta$ is the parameter of interest. Assume that the stochastic error term, $u_{it}$, takes the form

$$\begin{aligned} u_{it} &= \eta_i + \tau_t + \omega_{it}, \\ \omega_{it} &= \rho\omega_{it-1} + \upsilon_{it}. \end{aligned} \qquad (13)$$

The $\tau_t$ represent macroeconomic shocks captured by a series of time dummies, $\eta_i$ is an individual effect and $\upsilon_{it}$ is a serially uncorrelated mean zero error term. The other element of the error term, $\omega_{it}$, is allowed to have an AR(1) component (with coefficient $\rho$), which could be due to measurement error or slowly evolving technological change. Substituting equation (13) into equation (12) gives the dynamic equation

$$y_{it} = \pi_1 y_{it-1} + \pi_2 x_{it} + \pi_3 x_{it-1} + \eta_i^* + \tau_t^* + \upsilon_{it}. \qquad (14)$$

The common factor restriction (COMFAC) is $\pi_1\pi_2 = -\pi_3$. Note that $\tau_t^* = \tau_t - \rho\tau_{t-1}$ and $\eta_i^* = (1-\rho)\eta_i$.

In our main results section, we present several econometric estimates of production functions (random effects, within groups and GMM). The most rigorous approach follows that recommended by Blundell and Bond (2000), which uses a 'system GMM' approach to estimate equation (14) and then imposes the COMFAC restrictions by minimum distance. We now turn to describing the GMM approach in more detail.

How should equation (14) be estimated? If training is strictly exogenous and there are no dynamics (i.e. $\rho = 0$) then the only problem with ordinary least square (OLS) estimation of equation (12) is the presence of the

[12]One could argue that firm variables such as capital intensity and R&D should be excluded from the wage equation under competitive labour markets. However, these variables are typically quite informative in wage equations, either because they are picking up some measure of unobserved labour quality (Hellerstein and Neumark, 1999) or because of departures from perfect competition. In either case, omitting such variables is likely to cause bias on the training variable and our baseline specifications will include them.

individual effects, $\eta_i$. If these individual effects are uncorrelated with $x_{it}$ then the random-effects estimator is unbiased and efficient. If the individual effects are correlated with $x_{it}$ but remain strictly exogenous then although the random-effects estimator is biased, the within-groups estimator will be unbiased.

If we allow training to be endogenous (i.e. allowing training decisions to react to shocks to current productivity), we will require instrumental variables. In the absence of any obvious natural experiments, we consider moment conditions that will enable us to construct a GMM estimator for equation (14). A common method would be to take first differences of equation (14) to sweep out the fixed effects:

$$\Delta y_{it} = \pi_1 \Delta y_{it-1} + \pi_2 \Delta x_{it} + \pi_3 \Delta x_{it-1} + \Delta \tau_t^* + \Delta \upsilon_{it}. \tag{15}$$

As $\upsilon_{it}$ is serially uncorrelated, the moment condition

$$E(x_{it-2} \Delta \upsilon_{it}) = 0 \tag{16}$$

ensures that instruments dated $t - 2$ and earlier[13] are valid and can be used to construct a GMM estimator for equation (14) in first differences (Arellano and Bond, 1991). A problem with this estimator is that variables with a high degree of persistence over time (such as capital) will have very low correlation between their first difference ($\Delta x_{it}$) and the lagged levels being used as instruments (e.g. $x_{it-2}$). This problem of weak instruments can lead to substantial bias in finite samples.

Blundell and Bond (1998) point out that under a restriction on the initial conditions, another set of moment conditions are available:[14]

$$E(\Delta x_{it-1}(\eta_i + \upsilon_{it})) = 0. \tag{17}$$

This implies that lags of the first differences of the endogenous variables can be used to instrument the levels equation (14) directly. The econometric strategy is then to combine the instruments implied by the moment conditions (16) and (17). We stack the equations in differences and levels, i.e. equations (14) and (15). We can obtain consistent estimates of the coefficients and use these to recover the underlying structural parameters in equation (12).

The estimation strategy assumes the absence of serial correlation in the levels error terms ($\upsilon_{it}$).[15] We report serial correlation tests in addition to the

[13]Additional instruments dated $t - 3$, $t - 4$, etc. become available as the panel progresses through time.

[14]The restrictions are that the initial change in productivity is uncorrelated with the fixed effect $E(\Delta y_{i2}\eta_i) = 0$ and that initial changes in the endogenous variables are also uncorrelated with the fixed effect $E(\Delta x_{i2}\eta_i) = 0$.

[15]If the process is MA(1) instead of MA(0) then the moment conditions in equations (16) and (17) no longer hold. Nevertheless, $E(x_{it-3}\Delta\upsilon_{it}) = 0$ and $E(\Delta x_{it-2}(\eta_i + \upsilon_{it})) = 0$ remain valid, so earlier-dated lags could still be used as instruments. This is the situation empirically with the wage equations.

Sargan–Hansen test of the over-identifying restrictions in all the GMM results below.[16]

This GMM 'system' estimator has been found to perform well in Monte Carlo simulations (Blundell and Bond, 1998) and in the context of the estimation of production functions (Blundell and Bond, 2000). The procedure should also be a way of controlling for transitory measurement error (the fixed effects control for permanent measurement error). Random measurement error has been found to be a problem in the returns to human capital literature, typically generating attenuation bias (see Card, 1999).

In order to assess the importance of biases associated with fixed effects and endogeneity, we will estimate random effects, within groups and GMM estimates in section V.

Finally, consider two more issues which are harder to deal with: aggregation and training stocks vs. training flows. Estimation at the three-digit industry level has advantages but also disadvantages relative to micro-level estimation. The production function in equation (1) at the firm level describes the private impact of training on productivity. However, many authors, especially in the endogenous growth literature (e.g. Aghion and Howitt, 1998), have argued that there will be externalities to human capital acquisition. For example, workers with higher human capital are more likely to generate new ideas, which may spill over to other firms.[17] If spillovers are industry specific, this implies that there should be additional terms added to equation (5) representing training in other firms (e.g. the mean number of trained workers in the industry). In this case, the coefficient on training in an industry-level production function should exceed that in a firm-level production function.[18] Secondly, grouping by industry may smooth over some of the measurement error in the micro-data and therefore reduce attenuation bias.

On the negative side, there may be aggregation biases in industry-level data. *A priori* it is not possible to unambiguously sign these biases. We expect that the fixed effects will control for some of the problem. For example, we are taking logs of means and not the means of logs in aggregating equation (4), but so long as the higher-order moments of the distributions are constant over time in an industry then they will be captured by a fixed effect.[19] If the

[16]These are based on the first-differenced residuals, so we expect significant first-order serial correlation but require zero second-order serial correlation for the instruments to be valid. If there is significant second-order correlation, we need to drop the instruments back a further time period (this happens to be the case for the wage equation in the results below).

[17]Although there are many papers that examine externalities of R&D (e.g. see the survey by Griliches, 1992) and a few that look at education (Acemoglu and Angrist, 2000; Moretti, 2004), there are none that focus on training spillovers.

[18]For the same argument in the R&D context, see Griliches (1992).

[19]If they evolve at the same rate across industries, they will be picked up by the time dummies.

coefficients are not constant across firms in equation (4), but are actually random, this will also generate higher-order terms at the industry level. In the empirical results, we experiment with including higher-order moments and allowing the coefficients to vary across cross-sectional units.

Turning to the problem of training stocks and flows, note that the model in equation (1) assumes that we know the stocks of trained workers in an industry. What we actually have in the data are an estimate of the proportion of workers in an industry who received training in a given 4-week period (the training *flow*). Since individuals are sampled randomly over time in the LFS, this should be an unbiased estimate of the proportion of people in training in a given industry in a given year.[20] As an alternative to using the flow, we calculate a stock of training in an analogous way to using investment flows to calculate a capital stock through the perpetual inventory method (the main form of depreciation is the turnover rate). This is described in Dearden *et al.* (2005), Appendix A.

## IV. Data description

The database we construct combines several sources (see Dearden *et al.*, 2005, Appendix A for full details). The critical individual-level source is the individual-level UK LFS, which contained about 60,000 households per year. Most importantly, the LFS has a consistent training measure since 1984 as well as detailed information on skills, demographics, hours worked, tenure and wages. We work with this information aggregated by broadly three-digit industries. The LFS only started asking questions on wages at two points of time in 1997 (and at one point of time in 1992 when the panel was set up). We present some individual-level panel wage regressions at the end of section V for comparison.

The second major data set we use is the Annual Census of Production. This gives production statistics on capital, wages, labour and output, for industries in the production sector (manufacturing, mining and utilities). For the services industries, we drew on the OECD's ISDB data.

There was a change in Standard Industrial Classification (SIC) classification in 1992 which forced us to aggregate some of the industries and prevented us from using some of the industries after the change. Additionally, we insisted on having at least 25 individuals in each cell in each year. After matching the aggregated individual data from LFS, we were left with 94 industry groupings over (a maximum of) 14 years.

[20]If there are many multiple training spells in the month, we will underestimate the proportion of employees who are being trained. If Spring (the LFS quarter we use) is a particularly heavy training season then we will overestimate the proportion being trained in a year. These biases are likely to be small and offsetting.

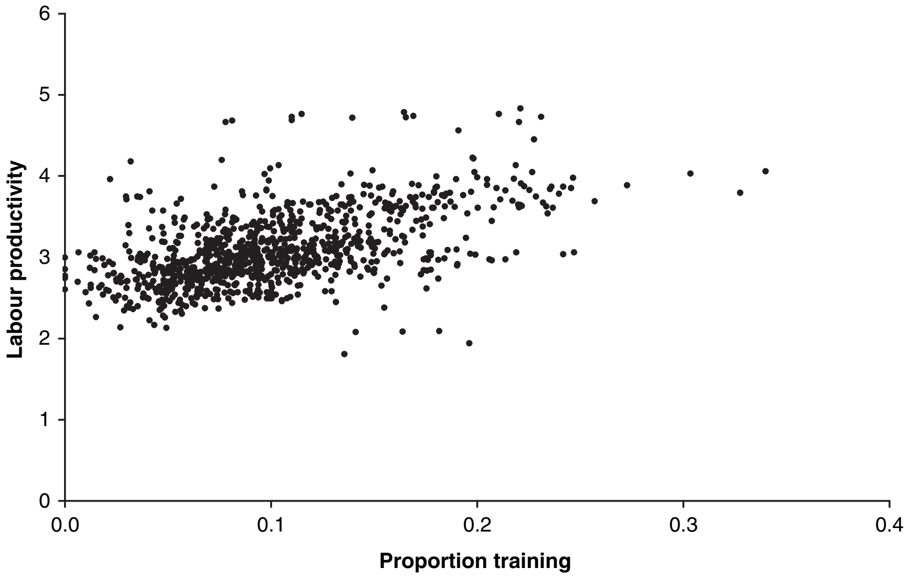


Figure 1. Labour productivity and training in British industries. Each point is an industry–year observation. The ordinary least square regression line has a slope of 4.91. Labour productivity is ln(value added per employee) aggregated from the Annual census of production and ISDB; training is the proportion of workers involved in training in the last 4 weeks from the LFS

The main LFS training question was '*over the 4 weeks ending Sunday . . . have you taken part in any education or training connected with your job, or a job that you might be able to do in the future?*'.[21] The average proportions of employees undertaking training grew steadily from about 8% in 1984 to 14% in 1990 where it stabilized for the next 6 years. Most of this growth was upgrading within industry rather than between industries.[22]

Figure 1 gives the scatterplot of labour productivity (log real value added per worker) against training propensity and Figure 2 repeats the exercise for log hourly wages. Not surprisingly, training has a strong positive correlation with both variables, but the association is somewhat weaker for wages than for productivity.

The outliers in both graphs tend to be in the service sector. Unfortunately, the published series for real value added and capital stocks are rather unreliable in the service sector. For example, in banking and financial services, measured real value added per person declined every year between 1983

[21]Unfortunately, it is not possible to separate out 'education' from 'training'.

[22]There is also a question on the length of the training spell, but this was only asked in particular years and there were too many missing values to use it as a separate regressor. Median spell length was 2 weeks and the mean was higher.

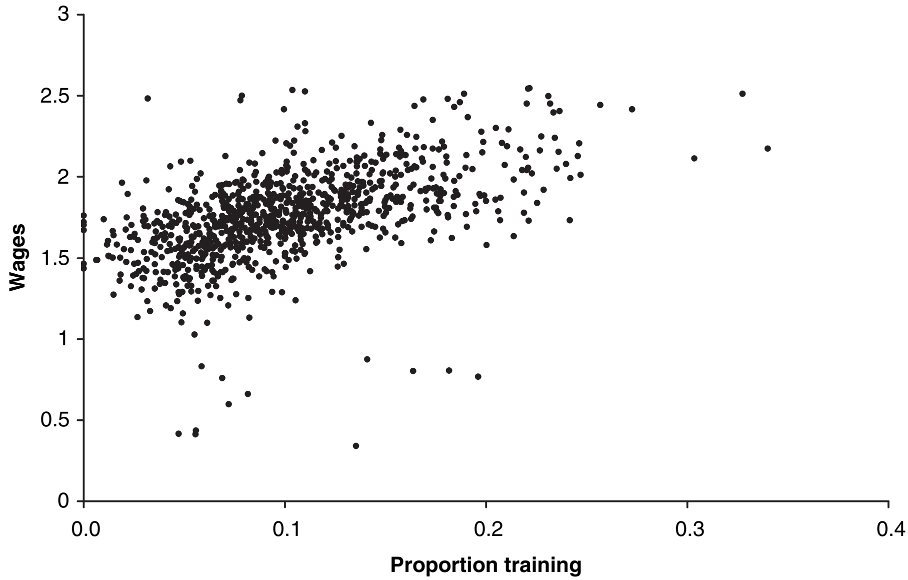


Figure 2. Wages and training in British industries. Each point is an industry–year observation. The ordinary least square regression line has a slope of 2.95. Wages are ln(hourly wages) from the ABI and ISDB (wages) and the Labour Force Survey (LFS) (hours); training is the proportion of workers involved in training in the last 4 weeks from the LFS

and 1996. Given the poor quality of the service sector production data, we reluctantly decided to focus the econometric part of the analysis on the production side of the economy. This is still a substantial share of the economy – about 50% of private sector net output in 1986.[23] Until robust measures of service sector productivity are developed, there is simply no alternative to the empirical strategy of focusing on the production sector.

Care must be taken in interpreting the scatterplots presented in Figures 1 and 2 as they say nothing about the causal impact of training on productivity or wages. High-training industries are characterized by higher fixed capital intensity, more professional workers, more educated workers and higher R&D (see Table A1 in Appendix A of Dearden *et al.*, 2005). We need to turn to an explicit econometric model to investigate whether there is a causal effect of training on productivity, and this forms the focus of the rest of the paper.

[23]This led to the loss of only 91 observations and the results are robust to including the service sector in the unweighted regressions. We generally weight the regressions by the number of LFS observations in order to reduce sampling variability. In the weighted regressions, including the service sector does have more substantial effects on the results because of its large employment shares. Full sets of these results are available on request from the authors.

## V. Results

### Baseline industry results

In Table 1, we present the basic results for the industry-level regressions treating all variables as exogenous. The first three columns have productivity (log real value added per head) as the dependent variable and the last three columns have wages as the dependent variable.

The first two columns are estimated by random effects; the only difference is that column (1) does not include the occupational controls. This omission makes some difference to the 'no qualifications' variable, which has a significantly negative association with productivity in column (1) but is insignificantly different from zero in column (2) – the occupational proportions (especially the professional/managerial category) do a better job at proxying for workforce skill than education.[24] The variables generally take their expected signs, although it is clear that there is some loss of precision when a full set of fixed effects is added in column (3). Capital per worker is strongly correlated with productivity, although the coefficient is lower (0.21–0.25) than capital's share of value added, which is about 30%. Worker turnover has a significantly negative association with productivity and R&D a significantly positive correlation. Younger workers (aged between 16 and 24) are significantly less productive than the 35- to 44-year-old group. Most importantly for our purposes, training has a statistically significant and economically important effect on productivity according to Table 1. The magnitude of the coefficient falls as we move to the more rigorous specification which controls for fixed effects, but the change is not dramatic. The estimates imply that raising the training variable by 1% point (say, from the 1996 economy-wide mean of about 14–15%) is associated with an increase in productivity of about 0.7%. We will return to the plausibility of the magnitude of these effects in the last subsection of section V.

The last three columns repeat the specifications but instead use ln(wages) as the dependent variable. The most interesting contrast for our purpose is the coefficient on training. As with productivity, training enters the earnings equation with a consistently positive and significant coefficient across all three columns. The magnitude of the coefficient is lower in the wage equation than in the productivity equation – about half the size. At face value, then, estimating the returns to training solely on the

[24]This conclusion does not change if we break down the qualifications into four groups. Machin, Vignoles and Galindo-Ruedo (2003) adopt a much finer classification of education using post-1992 LFS data where there are a larger number of observations. Exploiting the regional and industry aspects of the aggregated data, they find some role for college proportion, even in fixed-effect specifications.

TABLE 1

*Training, productivity and wages*

| | *(1)* | *(2)* | *(3)* | *(4)* | *(5)* | *(6)* |
|---|---|---|---|---|---|---|
| | *ln(value added per worker)* | | | *ln(wages)* | | |
| | *Random effects* | *Random effects* | *Within groups* | *Random effects* | *Random effects* | *Within groups* |
| Training | **0.788 (0.168)** | **0.700 (0.169)** | **0.696 (0.201)** | **0.425 (0.117)** | **0.344 (0.119)** | **0.365 (0.157)** |
| ln(capital/worker) | **0.252 (0.020)** | **0.244 (0.019)** | **0.212 (0.053)** | **0.058 (0.012)** | **0.051 (0.012)** | **0.069 (0.035)** |
| ln(hours/worker) | 0.184 (0.181) | 0.196 (0.181) | **0.275 (0.207)** | **0.274 (0.123)** | **0.272 (0.126)** | **0.310 (0.116)** |
| Lagged R&D intensity | **1.628 (0.430)** | **1.390 (0.428)** | **1.251 (0.662)** | −0.212 (0.284) | −0.356 (0.281) | −0.942 (0.717) |
| Worker turnover | −**0.632 (0.206)** | −**0.683 (0.207)** | −0.430 (0.332) | 0.132 (0.143) | 0.070 (0.145) | 0.163 (0.202) |
| Occupations: base group is manual workers | | | | | | |
| Managers | | **0.487 (0.123)** | **0.282 (0.131)** | | **0.324 (0.084)** | **0.195 (0.099)** |
| Clerical | | **0.366 (0.174)** | −0.076 (0.190) | | 0.161 (0.121) | −0.126 (0.121) |
| Personal/security | | −0.049 (0.355) | −0.522 (0.371) | | **0.504 (0.250)** | 0.204 (0.223) |
| Sales people | | 0.443 (0.276) | −0.078 (0.281) | | −0.037 (0.191) | −0.328 (0.190) |
| No qualifications | −**0.251 (0.096)** | −0.036 (0.109) | 0.107 (0.096) | −**0.145 (0.065)** | −0.033 (0.075) | 0.101 (0.069) |
| Experience: base group age 35–44 | | | | | | |
| Age 16–24 | −**0.579 (0.170)** | −**0.461 (0.172)** | −**0.390 (0.175)** | −**0.315 (0.118)** | −**0.259 (0.121)** | −0.153 (0.119) |
| Age 25–34 | −**0.341 (0.155)** | −0.282 (0.158) | −0.314 (0.171) | −0.198 (0.109) | −0.155 (0.110) | −0.196 (0.111) |
| Age 45–54 | −0.058 (0.158) | −0.042 (0.156) | −0.104 (0.160) | −0.139 (0.110) | −0.148 (0.111) | −0.150 (0.101) |
| Age 55–64 | 0.178 (0.190) | 0.244 (0.192) | 0.142 (0.237) | −**0.263 (0.133)** | −0.263 (0.136) | −**0.271 (0.138)** |
| Male | 0.037 (0.097) | 0.114 (0.099) | −0.116 (0.128) | **0.293 (0.064)** | **0.364 (0.065)** | −0.112 (0.078) |
| Small firm | 0.068 (0.112) | 0.016 (0.113) | 0.005 (0.127) | −0.118 (0.076) | −0.126 (0.076) | −0.056 (0.074) |
| Observations | 968 | 968 | 968 | 968 | 968 | 968 |
| Estimation period | 1984–96 | 1984–96 | 1984–96 | 1984–96 | 1984–96 | 1984–96 |

*Notes*: Standard errors (robust to heteroskedasticity) are given in parentheses under coefficients. In the first three columns the dependent variable is ln(value added per worker) and in the last three columns the dependent variable is ln(wages). Bold typeface indicates that the variable is significant at the 5% level. All regressions include a full set of regional dummies (10), time dummies (12) and tenure dummies (6). Observations are weighted by number of individuals in a Labour Force Survey industry cell. Random effects are estimated by generalized least squares. Within groups are estimated by least squares dummy variables (85 industries).

basis of wage equations would generate an underestimate of the importance of work-related training.[25]

Turning to the other variables in the wage equation, the signs of most of them are the same as those in the productivity equations, although there are some differences. As expected, earnings are significantly higher in more capital-intensive, hours-intensive and highly skilled industries. The R&D coefficient is surprisingly negative (although insignificantly different from zero), but this turns out to be because of mis-specified dynamics – including longer lags of R&D demonstrates there is actually a positive correlation of technology with wages.[26]

An important concern with Table 1 is that we do not allow for the endogeneity of training or other suspected endogenous variables. To deal with this, we implemented the GMM approach described in section III above. Table 2 contains a summary of the main results.[27] All the same variables are included in these regressions as in Table 1, but we report only the key coefficients to preserve space.

In column (1) we present the production function and in column (2) we present the wage equation. The GMM estimates tell a similar story to the within-groups estimates. Training has a positive and significant impact on both productivity and wages, although the training coefficient in the production function remains almost twice the size of the coefficient in the wage equation (0.60 vs. 0.35). There are some minor changes to the other coefficients – the coefficient on capital intensity has risen to 0.33 in the production function, the R&D coefficient is positively signed in the wage equation and the coefficient on hours is somewhat larger in magnitude than in Table 1.

The diagnostics reported at the base of the table are also satisfactory – there is no sign of second-order serial correlation (in the first-differences residuals) and the Sargan test of over-identifying restrictions does not reject. Note that the wage regression uses instruments dated $t - 3$ and before in the differenced equation (and dated $t - 2$ in the levels equation). This is because there were some signs of significant second-order serial correlation using $t - 2$ dated instruments in the wage equation, which invalidates the Instrumental Variables (IVs) (we dropped one period in

[25]A test of the equality of the coefficients on training in the wage and productivity equations rejected equality at the 0.10 level for training ($P$-value = 0.068). We would, however, expect the coefficient on training in the wage equation ($\lambda - 1$) to be lower than in the production function as the training coefficient in equation (7) is $\alpha(\gamma - 1)$. Although a test of the equality between $\lambda$ and $\gamma$ cannot be rejected at the 0.05 level for the sample as a whole it *can* be rejected for the 'low wage' industries – see footnote 33 below.

[26]Consistent with the findings of, *inter alia*, Bartel and Sicherman (1999).

[27]Table B2 in Appendix B of Dearden *et al.* (2005), has more detailed results, and even more detailed specifications are available from the authors or in Dearden, Reed and Van Reenen (2000).

TABLE 2

*Production functions and wage equations estimated by general method of moments (GMM)*

| | *(1)* *ln(real value added per worker)* | *(2)* *ln(wages)* |
|---|---|---|
| Training | **0.602** (**0.181**) | **0.351** (**0.074**) |
| ln(capital/worker) | **0.327** (**0.016**) | **0.106** (**0.011**) |
| ln(hours/worker) | **0.498** (**0.064**) | **0.489** (**0.027**) |
| Lagged R&D intensity | **1.905** (**0.262**) | **0.443** (**0.182**) |
| Proportion of employees who are professionals or managers | **0.306** (**0.068**) | **0.160** (**0.034**) |
| Autocorrelation coefficient ($\rho$) | **0.741** (**0.015**) | **0.797** (**0.013**) |
| LM1 (d.f.) | −4.892 (85) | −6.053 (85) |
| [*P*-value] | [0.00] | [0.00] |
| LM2 (d.f.) | −0.940 (85) | −1.44 (85) |
| [*P*-value] | [0.347] | [0.158] |
| Sargan (d.f.) | 8.819 (121) | 11.83 (146) |
| Instruments | $(\text{TRAIN})_{t-2,t-3}$, $\ln(Q/N)_{t-2,t-3}$, $\ln(\text{Hrs}/N)_{t-2,t-3}$, $\ln(K/N)_{t-2,t-3}$ in differenced equations; $\Delta(\text{TRAIN})_{t-1}$, $\Delta\ln(\text{Hrs}/N)_{t-1}$, $\Delta\ln(K/N)_{t-1}$ in levels equations | $(\text{TRAIN})_{t-3,..,t-5}$, $\ln(Q/N)_{t-3,..,t-5}$, $\ln(\text{Hrs}/N)_{t-3,..,t-5}$, $\ln(K/N)_{t-3,..,t-5}$ in differenced equations; $\Delta(\text{TRAIN})_{t-2}$, $\Delta\ln(\text{Hrs}/N)_{t-2}$, $\Delta\ln(K/N)_{t-2}$ in levels equations |
| Estimation period | 1984–96 | 1985–96 |
| Observations | 898 | 883 |

*Notes*: Estimation by GMM-SYS in Arellano and Bond (1998) DPD-98 package written in GAUSS; one-step robust estimates reported. All regressions include the current values of all the variables in columns (3) and (6) of Table 1 (i.e. worker turnover, occupational proportions, qualifications, age, tenure, gender, region, firm size and time dummies). Capital intensity, training, hours and lagged productivity are always treated as endogenous. The other variables are assumed exogenous. One-step SEs (robust to arbitrary heteroskedasticity and autocorrelation of unknown form) are given in parentheses under coefficients (variables significant at 5% level are in bold). LM1 (LM2) is a Lagrange Multiplier test of first- (second)-order serial correlation distributed $N[1, 0]$ under the null (see Arellano and Bond, 1991). Sargan is a chi-squared test of the over-identifying restrictions. Observations are weighted by number of individuals in a Labour Force Survey industry cell. Full details in Table B2 in Appendix B of Dearden *et al.* (2005).

order to be able to use the longer lags in estimation). Using the (invalid) instruments on the longer time period gave a coefficient (SE) on training of 0.141 (0.067) in the wage equation.[28]

[28]Using the shorter time period with longer-dated instruments in the production function gave a coefficient (SE) on training of 1.043 (0.325). See Table B2 in Appendix B of Dearden *et al.* (2005) for full details.

TABLE 3

*Robustness tests of the production function*

| *Row* | *Robustness test* | *Observations* | *Training coefficient (SE)* |
|---|---|---|---|
| 1 | Original training coefficient in production sector, Table 1 column (3) | 968 | **0.696 (0.201)** |
| 2 | Using 'stock' of trained workers instead of flows | 968 | **0.775 (0.189)** |
| 3 | Using the balanced panel only | 572 | 0.508 (0.289) |
| 4 | Conditioning on wage in productivity regression (to control for any residual unobserved worker quality) | 968 | Training: **0.659 (0.219)** Wage coeff.: **0.099 (0.130)** |
| 5 | Including service sectors | 1,059 | **0.727 (0.206)** |
| 6 | Include union density (only available 1989–96) | 547 | Training: **0.604 (0.266)** Union: −0.177 (0.183) |
| 7 | Allow all industries to have different training coefficients | 968 | Mean of heterogeneous coefficients: 0.505 |
| 8 | Allow non-constant returns | 968 | **0.725 (0.201)** |
| 9 | Estimating a translog production function | 968 | **0.703 (0.201)** |
| 10 | Estimation by non-linear least squares | 968 | **0.518 (0.197)** |
| 11 | Estimation on 1993–2001 data (region–industry cells) | 1,873 | **0.436 (0.188)** |

*Notes*: These all use the specification in column (3) of Table 1 (unless otherwise specified). Estimation by within groups; robust SEs in parentheses (except row 10). Bold typeface indicates that the variable is significant at the 5% level.

### Robustness of the results

We conducted a large number of robustness tests on the models in Tables 1 and 2. Table 3 reports some of these. Given the similarity of the within-groups and GMM results, we performed these tests on the within-groups specifications of Table 1, column (3). The first row of Table 3 simply reports the coefficient and SE from that column. Using the stock of trained employees instead of the flow (calculated allowing for depreciation due to inter-firm turnover) results in very similar results in row 2.[29] Keeping only industries

[29]The non-fixed-effects results were significantly different, but the deviations around the fixed effect in the training stock are dominated by the flow, explaining the similarity of the results. In addition to using the empirical worker turnover rates, we assumed an exogenous depreciation of training at 40% per annum (see Appendix A of Dearden *et al.*, 2005). The coefficient was stable to reasonable changes in these parameters (e.g. increasing the depreciation rate to 50% p.a. increased the coefficient to 0.81, to 60% to 0.82; decreasing the depreciation rate to 30% reduced the coefficient to 0.70).

that had 14 continuous years of data (balanced panel) in row 3 means losing 40% of the observations; the coefficient falls, but the change is not significant (cf. Nickell, 1981). The fourth row includes average wages on the right-hand side of the production function as a measure of unobserved worker quality; although wages have a positive coefficient, the training association remains robust. In row 5 we include all the service sectors, ignoring our concerns over data quality. The coefficient on training rises to 0.73 and remains significant.

As training is correlated with unionization, we could be picking up 'collective voice' effects in the main results. Union membership is only available in LFS since 1989. Despite the loss in sample in row 6, the training effect is robust to inclusion of union density (density is insignificantly negatively associated with productivity). In row 7 we allow the training effect to be different in each of the 85 industries; the mean of these heterogeneous coefficients is close to the pooled results. The next two rows allow for more general functional forms, first relaxing constant returns (row 8) and then estimating a translog production function (row 9); in both cases, the training coefficient is essentially unchanged. Row 10 gives the results from a non-linear least squares estimation of equation (8) again showing no significant difference.

We also compared our results with a recent paper (Machin *et al.*, 2003) which has built up similar data to our own covering a more recent period and exploiting the larger size of the LFS post-1992 to construct industry-by-region cells. Against these advantages, their data set has a shorter time-series component (1993–2001) and lacks some of the covariates we use. Re-estimating identical specifications on their data set gives an estimate of the training association with productivity of 0.436 with a SE of 0.188 (see row 11 of Table 3). This is lower, but is still significant and remains well within two SEs of our main results.[30] On our data set, we tested whether there was a tendency for the training coefficient to fall (or rise) over time in the production function, but we found it to be stable.[31]

Does the 'wedge' between the wage and productivity effect of training arise from specific human capital or imperfect competition? Under most forms of imperfect competition, we conjectured that the wedge would be larger in those industries where workers were earning less than would be implied by their human capital (i.e. inter-industry wage premiums were low). This could be because the 'low-paying' industries were monopsonistic with large search

[30]The specification is identical to column (3) of Table 1 except we drop the occupational proportions and R&D and include employment. On our data, this gives a coefficient (SE) on TRAIN of 0.732 (0.205).

[31]For example, interacting TRAIN with a trend in the production function gave a coefficient of 0.003 with a SE of 0.044.

frictions or because workers are more able to capture the quasi-rents from training in the 'high-paying' industries.

In order to identify such industries, we used estimates of inter-industry wage premiums taken from the US Current Population Survey (CPS).[32] We matched the US industries to the UK industries and split the sample at the median sectoral wage premium. Allowing an interaction between training and this industry split revealed that the wedge between the training effect on productivity and the training effect on wages was solely within the 'low-wage' industries. To be precise, including an interaction in the wage equation between training and low-wage industries gave a coefficient (SE) of −0.664 (0.196) on the interaction and 0.531 (0.113) on the linear training effect. In the production function, the interaction was 0.332 (0.297) – positive but insignificant (the linear training term took a coefficient of 0.612 with an SE of 0.172).[33] In other words, in the 'high wage premium' industries, there was no significant difference between the impact of training on productivity and the impact of training on wages. The fact that our results are driven by the wedge in low-paying sectors is tentative evidence in favour of a monopsony/search interpretation.

This evidence is open to the critique that firm-specific training may be systematically more prevalent in the low-wage sectors (although a priori the usual view is that 'good jobs' are more likely to have more specific skills). There are several questions in LFS that could be interpreted as general vs. specific training, so we used them to see if the coefficients differed significantly with training type – they did not. For example, there are questions related to off-the-job training (more general) and on-the-job training (more specific). The proportion of off-the-job training produced a coefficient (SE) of 0.005 (0.018) when added to the wage regression and a coefficient (SE) of 0.018 (0.029) when added to the production function. We view this not as any rejection of specific human capital theory *per se*, but rather as an indication the type of human capital is intrinsically difficult to measure. Furthermore, the LFS questions are not asked in all years and have many missing values.

[32]Estimating inter-industry wage premiums from UK wages would have been more problematic as these could reflect endogenous influences – US wage-setting will be driven by the structural characteristics of the industries in question. These US inter-industry wage premiums were generated from individual-level wage regressions from the 1986 CPS merged outgoing rotation files. The wage regressions included years of schooling, a quartic in experience, gender, marital status, gender × marriage interactions, race, Standard Metropolitan Statistical Area and regional dummies. The data were kindly provided by Steve Pischke (see Acemoglu and Pischke, 2003, for details).

[33]A test of the equality between the effects of training on wages ($\lambda$) and on productivity ($\gamma$) can be rejected at the 0.05 level for the 'low wage' industries ($P$-value = 0.001), but cannot be rejected for the 'high wage' industries ($P$-value = 0.752).

**Quantifying the effects of training**

Our key qualitative conclusions are: first, there is a significant impact of training on productivity; and, secondly, the effects of training on productivity are larger than the effects of training on wages. But, quantitatively, how economically significant is the magnitude of the training effect?

Interpreting the exact magnitude of the coefficients is difficult, but the implied effects are large. From Tables 1 and 2, we conservatively take the coefficient on training in the productivity regressions to be about 0.6 and the coefficient on training in the wage regressions to be about 0.3. This would imply that a 10% point increase in the training measure is associated with a 6% increase in productivity and a 3% increase in wages.

Relative to the returns-to-schooling literature, the training impacts appear high.[34] Card (1999) puts the impact of a year of schooling on wages at about 10%, so our baseline impact of 0.3 is about three times as large. Given that the typical time in training during the 4-week period is under a month (the median is 2 weeks, the mean is higher), the returns to a month of training appear even more impressive. For example, an increase in our key variable, TRAIN, of 10% would imply a typical worker only spent 5% extra of his time in training, if training spells were on average 2 weeks long.

Of course, there may be remaining econometric problems we have not controlled for generating this difference. But assuming the training effect is not a statistical artefact, there remain at least two possible explanations for the training coefficients being larger than conventional estimates of the return to schooling. First, work-related training may have a higher private return than schooling as training is more directed at raising productivity in employment. Training is also likely to have a faster rate of depreciation than schooling, so it requires a higher year-on-year return in order to give incentives for investment.[35] Secondly, there may be externalities associated with training that are missed in the conventional schooling literature, which focuses on private returns whereas we look at returns to the industry as a whole (cf. Moretti, 2004).

To investigate the externality issue, we estimated some individual-level wage regressions on the LFS panel. If the private returns to training are higher than the social returns, we might expect to see a similarly high coefficient in the individual-level wage regression. We used the individual-level equivalents of the variables in the industry-level regressions. To construct the proportion of the year spent in training, we used the LFS panel which follows individuals

[34]Compared with existing UK estimates of the training effects on wages (e.g. Booth, 1993; Blundell *et al.*, 1996), our estimates are actually lower (see Dearden *et al.*, 2000, for a detailed comparison).

[35]See Heckman, Lochner and Todd (2003) for a recent discussion of interpretation of the schooling coefficient in wage regressions.

for five quarters and asks individuals the training question in each quarter. We defined a dummy variable (TRAIND4) indicating whether the individual had been involved in some training in *all* of the previous four quarters. We also defined dummies for if the individual had been in training for three quarters (TRAIND3), two quarters ( TRAIND2), one quarter (TRAIND1) or not at all (TRAIND0). Using TRAIND0 as the omitted base, the results we obtained from a typical regression were:[36]

$$\begin{aligned}\ln(\text{wage}) = {} & 0.165(0.033)\text{TRAIND4} + 0.092(0.023)\text{TRAIND3} \\ & + 0.125(0.019)\text{TRAIND2} + 0.078(0.015)\text{TRAIND1} + \text{controls}.\end{aligned}$$

Longer lengths of time in training are associated with significantly higher wages.[37] The coefficient on receiving training in all four quarters was 0.165; this is comparable with the industry-level coefficient of 0.350. Taken literally, this would suggest that about half of the impact of training on wages at the industry level is attributable to externalities.

If we include a set of industry dummies (which will include potential spillovers), the coefficient on TRAIND4 falls from 0.16 to 0.13. If we also include the initial wage in the first quarter (to control for unobserved heterogeneity), the coefficient falls even further to 0.079. So these impacts of a 'year' of training are rather similar to the conventional impacts of the returns to a year in school.

Our conclusion from this exercise is that the larger magnitude of the training effects in this paper primarily reflects our strategy of estimating at a level above the individual worker. This was forced upon us by the absence of adequate data on firm productivity and training, but also because of our desire to incorporate externalities. The results are therefore consistent with a story that stresses externalities to training.

Even if there remain econometric problems that have caused us to overestimate the impact of training at the industry level, it is hard to see why this would not also bias upwards the training coefficient in the production function and wage equation to a similar extent. Therefore, even if one disputes the exact quantitative magnitude of the training effect, our key qualitative conclusion that the productivity impact of training is greater than the wage impact should still be valid (this is also a feature of the firm-level results in Appendix B of Dearden *et al.*, 2005).

[36]Estimation was by OLS; robust SEs are given in parentheses. Controls include gender, age, areas (20), employer size, occupational dummies (8), no qualification dummy, and a dummy for part-time status. Results are for the production sector only. The quarterly LFS panel 1997–98 was used as two wage observations per individual did not exist in the LFS prior to this. There were 3,998 observations. Full results are available on request from the authors.

[37]The training effects are not monotonic. There is even a perverse fall in the coefficient on being in training three relative to two quarters, although the coefficients are not significantly different.

## VI. Conclusions

In this paper, we have examined the issue of the impact of private sector training on productivity. Rather than simply use wages as a measure of productivity, we have presented (for the first time) estimates of the impact of training on productivity over a long time period. We have assembled a dataset that aggregates individual-level data on training and establishment data on productivity and investment into an industry panel covering 1983–96. We controlled for unobserved heterogeneity and the potential endogeneity of training using a variety of methods including GMM system estimation.

Using these new data, we have identified a statistically and economically significant effect of training on productivity in the UK. An increase of 1% point in the proportion of employees trained is associated with about a 0.6% increase in productivity and a 0.3% increase in wages. The impact of training on productivity is robust to a large number of robustness tests.

We argued that the methodologies in the existing literature may underestimate the importance of training. The focus on wages as the only relevant measure of productivity ignores the additional productivity benefits the firm may capture. The coefficient of training in the production function was around twice as large as the coefficient in the wage equation. This result could occur even under standard specific human capital theory. But it could also arise for a number of other reasons due to imperfect competition in the labour market (and we have presented some evidence consistent with this hypothesis). Clearly, further research is needed to distinguish between these theories.

Finally, a comparison between the industry- and individual-level wage regressions suggests that our industry-level analysis may capture externalities from training that are missed out in the micro-level studies. An important avenue of future research would include probing the returns to training by combining enterprise data with industry-level data to investigate the externalities story in greater detail.

*Final Manuscript Received: September 2005*

## References

Acemoglu, D. and Angrist, J. (2000). 'How large are the social returns to education? Evidence from compulsory schooling laws', *NBER Macroeconomics Annual*, Vol. 15, pp. 9–59.

Acemoglu, D. and Pischke, S. (1999). 'The structure of wages and investment in general training', *Journal of Political Economy*, Vol. 107, pp. 539–572.

Acemoglu, D. and Pischke, S. (2003). 'Minimum wages and on the job training', *Research in Labor Economics*, Vol. 22, pp. 159–202.

Aghion, P. and Howitt, P. (1998). *Endogenous Growth Theory*, MIT Press, Cambridge, MA.

Arellano, M. and Bond, S. (1991). 'Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations', *Review of Economic Studies*, Vol. 58, pp. 277–297.

Arellano, M. and Bond, S. (1998). *Dynamic Panel Data Estimation Using DPD for GAUSS*, Mimeo, Institute for Fiscal Studies, London.

Ballot, G. and Taymaz, E. (1998). *Firm Sponsored Training and Performance. A Comparison Between French and Swedish Firms*, Mimeo, ERMES, Paris.

Ballot, G., Fakhfakh, F. and Taymaz, E. (1998). 'Formation continue, recherche et développement, et performance des entreprises', *Formation Emploi*, Vol. 64, pp. 43–55.

Ballot, G., Fakhfakh, F. and Taymaz, E. (2002). *Who Benefits from Training and R&D: the Firm or the Workers?*, Economic Research Center Working Papers in Economics 02/01, Middle East Technical University, Ankara, Turkey.

Barron, J. M., Black, D. A. and Loewenstein, M. A. (1989). 'Job matching and on-the-job training', *Journal of Labour Economics*, Vol. 7, pp. 1–19.

Bartel, A. P. (1994). 'Productivity gains from the implementation of employee training programmes', *Industrial Relations*, Vol. 33, pp. 411–425.

Bartel, A. P. (1995). 'Training, wage growth, and job performance: evidence from a company database', *Journal of Labour Economics*, Vol. 13, pp. 401–425.

Bartel, A. and Sicherman, N. (1999). 'Technological change and wages: an inter-industry analysis', *Journal of Political Economy*, Vol. 107, pp. 285–325.

Barrett, A. and O'Connell, P. (2001). 'Does training generally work? The returns to in-company training', *Industrial and Labor Relations Review*, Vol. 54, pp. 647–683.

Black, S. E. and Lynch, L. M. (1996). 'Human-capital investments and productivity', *American Economic Review*, Vol. 86, pp. 263–267.

Black, S. E. and Lynch, L. M. (2001). 'How to compete: the impact of workplace practices and information technology on productivity', *Review of Economics and Statistics*, Vol. 83, pp. 434–445.

Blanchflower, D. and Lynch, L. M. (1992). *Training at Work: A Comparison of US and British Youths*, Discussion Paper No. 78, Centre for Economic Performance, London School of Economics, London.

Blundell, R. and Bond, S. (1998). 'Initial conditions and moment restrictions in dynamic panel data models', *Journal of Econometrics*, Vol. 87, pp. 115–143.

Blundell, R. and Bond, S. (2000). 'GMM estimation with persistent panel data: an application to production functions', *Econometric Reviews*, Vol. 19, pp. 321–340.

Blundell, R., Dearden, L. and Meghir, C. (1996). *The Determinants of Work-related Training in Britain*, Institute for Fiscal Studies, London.

Boon, M. and van der Eijken, B. (1997). *Employee Training and Productivity in Dutch Manufacturing Firms*, Research Paper No. 9716, Department of Statistical Methods, Statistics Netherlands, Voorburg.

Booth, A. (1991). 'Job-related formal training: who receives it and what is it worth?', *Oxford Bulletin of Economics and Statistics*, Vol. 53, pp. 281–294.

Booth, A. (1993). 'Private sector training and graduate earnings', *Review of Economics and Statistics*, Vol. 75, pp. 164–170.

Booth, A., Francesconi, M. and Zoega, G. (1999). 'Training, rent sharing and unions', Mimeo, University of Essex, Essex.

Card, D. (1999). 'The causal effect of education on earnings', in Ashenfelter O. and Card D. (eds), *Handbook of Labour Economics*, Elsevier, Amsterdam, pp. 1801–1863.

Carriou, Y. and Jeger, F. (1997). 'La formation continue dans les entreprises et son retour sur investissement', *Economie et Statistique*, Vol. 303, pp. 45–58.

Dearden, L., Reed, H. and Van Reenen, J. (2000). *Who Gains when Workers Train? Training and Corporate Productivity in a Panel of British Industries*, Working Paper No. 00/04, Institute for Fiscal Studies.

Dearden, L., Reed, H. and Van Reenen, J. (2005). *The Impact of Training on Productivity and Wages: Evidence from British Panel Data*, IFS Working Papers No. W05/12, Institute for Fiscal Studies, August.

Delame, E. and Kramarz, F. (1997). 'Entreprises et formation continue', *Economie et Prevision*, Vol. 127, pp. 63–82.

Green, F. and Steedman, H. (1997). *Into the Twenty First Century: An Assessment of British Skill Profiles and Prospects*, Special Report, Centre for Economic Performance, London School of Economics, London.

Greenhalgh, C. (2002). 'Adult vocational training in France and Britain', *Fiscal Studies*, Vol. 23, pp. 223–263.

Greenhalgh, C. and Stewart, M. (1987). 'The effects and determinants of training', *Oxford Bulletin of Economics and Statistics*, Vol. 49, pp. 171–189.

Griliches, Z. (1992). 'The search for R&D spillovers', *Scandinavian Journal of Economics*, Vol. 94, pp. S29–S47.

Grunfeld, D. and Griliches, Z. (1960). 'Is aggregation necessarily bad?', *Review of Economics and Statistics*, Vol. 42, pp. 1–13.

Hægeland, T. and Klette, T. (1999). 'Do higher wages reflect higher productivity? Education, gender and experience premiums in a matched plant-worker data set', in Haltiwanger J., Lane J., Spletzer J. R., Theeuwes J. and Troske K. (eds), *The Creation and Analysis of Employer–Employee Matched Data*, North Holland, Amsterdam.

Haskel, J., Hawkes, D. and Pereira, S. (2003). *Skills, Productivity in the UK Using Matched Establishment, Worker and Workforce Data*, Ceriba, Discussion Paper, May.

Heckman, J., Lochner, L. and Todd, P. (2003). *Fifty Years of Mincer Earnings Regressions*, National Bureau of Economic Research, Working Paper No. 9732.

Hellerstein, J. and Neumark, D. (1999). 'Sex, wages and productivity: an empirical analysis of Israeli firm level data', *International Economic Review*, Vol. 40, pp. 95–123.

Hellerstein, J., Neumark, D. and Troske, K. (1999). 'Wages, productivity and worker characteristics: evidence from plant level production functions and wage equations', *Journal of Labor Economics*, Vol. 17, pp. 409–446.

Holzer, H. (1990). 'The determinants of employee productivity and earnings', *Industrial Relations*, Vol. 29, pp. 403–422.

Holzer, H., Block, R., Cheatham, M. and Knott, J. (1993). 'Are training subsidies for firms effective?', *Industrial and Labor Relations Review*, Vol. 46, pp. 625–636.

Ichniowski, C., Shaw, K. and Prennushi, G. (1997). 'The effects of human resource management practices on productivity', *American Economic Review*, Vol. 87, pp. 291–313.

Jones, P. (2001). 'Are educated workers really more productive?', *Journal of Development Economics*, Vol. 64, pp. 57–79.

de Koning, J. (1994). 'Evaluating training at the company level', in McNabb R. and Whitfield K. (eds), *The Market for Training*, Avebury, Aldershot, pp. 87–105.

Krueger, A. and Rouse, C. (1998). 'The effects of workplace education on earnings, turnover and job performance', *Journal of Labor Economics*, Vol. 16, pp. 61–94.

Lazear, E. (1979). 'Why is there mandatory retirement?', *Journal of Political Economy*, Vol. 87, pp. 1261–1284.

Lillard, L. A. and Tan, H. W. (1992). 'Private sector training: who gets it and what are its effects?', *Research in Labour Economics*, Vol. 13, pp. 1–62.

Lynch, L. M. (1992). 'Private sector training and the earnings of young workers', *American Economic Review*, Vol. 82, pp. 299–312.

Machin, S., Vignoles, A. and Galindo-Rueda, F. (2003). *Sectoral and Area Analysis of the Economic Effects of Qualifications and Basic Skills*, Research Report No. RR465, Department for Education and Skills, London.

Moretti, E. (2004). 'Workers' education, spillovers and productivity: evidence from Plant-Level Production Functions', *American Economic Review*, Vol. 94, pp. 656–690.

National Skills Task Force (1998). *Towards a National Skills Agenda*, Department for Education and Employment, Sudbury.

Nickell, S. J. (1981). 'Biases in dynamic models with fixed effects', *Econometrica*, Vol. 49, pp. 1417–1426.

O'Mahony, M. (1998). 'Anglo-German productivity differences: the role of broad capital', *Bulletin of Economic Research*, Vol. 50, pp. 19–36.

Winkelmann, R. (1994). *Training, Earnings and Mobility in Germany*, Discussion Paper No. 982, Centre for Economic Policy Research.