

A GENERAL EMPIRICAL MODEL TO ESTIMATE THE IMPACT OF TRAINING ON INDIVIDUAL LABOUR MARKET PERFORMANCE

BACKGROUND MATERIAL FOR THE 2004 EDITION OF THE OECD EMPLOYMENT OUTLOOK

The general empirical model used in Chapter 4 of OECD (2004) can be considered an extension of that proposed by Loewenstein and Spletzer (1998). Denote with V_{ijt} the value for the worker i of a job match with the firm j at time t . In the simplest case, this value can be seen as the stream of expected revenues that the worker i can obtain from being employed in firm j at time t . In a narrow sense, we can think of this value as the current wage. However, more generally, this value may include the worker's valuation of its employment security.

Whatever its precise definition, which will depend on the specific empirical application, let us assume that it can be written as

$$V_{ijt} = \beta \mathbf{X}_{it} + \gamma T_{ijt}^c + \delta T_{ijt}^p + y_t + \mu_i + v_{ijt} + \varepsilon_{it} \quad [1]$$

where \mathbf{X}_{it} is a vector of time-varying individual characteristics, T_{ijt}^c and T_{ijt}^p are the stock of training with current and previous employers, respectively, while y_t , μ_i , v_{ijt} are year (or country per year) effects, individual fixed effects and job-match-specific effects (with v_{ijt} taking the value v_{ij} if the worker i has a job with firm j at time t and 0 otherwise), respectively, and ε_{it} is a standard independently distributed random disturbance.

The inclusion of an individual fixed effect allows identifying the coefficient of all stock variables (such as training) for which only changes within the sample period are observable (depreciation is ruled out for convenience). Assuming that [1] is valid is equivalent to ruling out time-variant heterogeneity, which is not due to observable characteristics (such as the training stock), the job-match or serially uncorrelated random disturbances. However, if match-specific effects are included in the empirical specification, since T_{ijt}^p is invariant within each specific job-match, its impact cannot be identified.

Loewenstein and Spletzer (1998) show that if $\gamma < \delta$ then estimating [1] by omitting match specific effects (but including individual fixed effects) would yield an estimate $\hat{\delta} < \delta$, provided that dummies for the number of job changes are included in the specification. Equivalently, the same result can be obtained by estimating model [1] in first differences using OLS, omitting match-specific effects and including dummies for job changes (see Baltagi, 1995, on the equivalence between the fixed effect

estimator applied to a level equation and the OLS estimator applied to a first-differenced equation). To see Loewenstein and Spletzer's result in a simple way, consider the following transformation of [1] for the case in which the individual separates from firm k to take a job in firm j :

$$V_{ijt} - V_{ikt-1} = \beta \Delta \mathbf{X}_{it} + \gamma \tau_{ijt} + (\delta - \gamma) T_{ikt-1}^c + \Delta y_t + u_{ijkt} \quad [2]$$

where τ_{ijt} stands for the flow of training received in the last period (after the job change), $u_{ijkt} = (V_{ijt} - V_{ikt-1}) + \Delta \varepsilon_{it}$ and Δ stands for the first difference operator. If [1] is the true model, T_{ikt-1}^c is likely to be correlated with the error term u_{ijkt} , since it is likely to be correlated with $(V_{ijt} - V_{ikt-1})$. However, a worker will quit *voluntarily* firm k for firm j at time $t-1$ if and only if $E(V_{ijt-1} - V_{ikt-1}) > 0$. This implies that the worker will quit if and only if

$$(\delta - \gamma) T_{ikt-1}^c + (V_{ijt} - V_{ikt-1}) = (\delta - \gamma) T_{ikt-1}^c + (V_{ij} - V_{ik}) > 0. \quad [3]$$

If $\gamma < \delta$ then [3] implies that the greater the value of T_{ikt-1}^c the smaller the minimum value of $(V_{ijt} - V_{ikt-1})$ that will suffice to induce a worker to switch jobs; therefore the sign of the correlation between T_{ikt-1}^c and $(V_{ijt} - V_{ikt-1})$ will be negative. If the worker has been laid off by firm k he/she will not compare V_{ij} with V_{ik} ; therefore $\text{corr}(T_{ikt-1}^c, (V_{ijt} - V_{ikt-1})) = -\text{corr}(T_{ikt-1}^c, V_{ikt-1})$, which is likely to be negative independently of the sign of $(\delta - \gamma)$. Estimating [2] with OLS will therefore yield a downward biased estimate of $(\delta - \gamma)$ if $\gamma < \delta$ — that is an estimate biased against finding a transferable positive impact of training.

As regards obtaining an unbiased estimate of γ , an alternative to including match-specific effects is to subtract job-match-specific means from the stock of training taken with the current employer, since $\text{corr}((T_{ijt}^c - \bar{T}_{ij}^c), V_{ij}) = 0$ by construction. This is the methodology used in the wage regressions in Chapter 4 of OECD (2004). A sensitivity analysis was also undertaken by estimating [1] directly, including match-specific effects, and revealed that the two procedures give extremely close results as regards training taken with the current employer.

The procedure discussed above is nowadays standard for the analysis of the relationship between training and wages (see Loewenstein and Spletzer, 1998, 1999; Parent, 1999; Booth and Bryan, 2002; and Gerfin, 2003). Nevertheless, one could use the same approach based on [1] also to estimate the effect of training on labour force participation and/or unemployment status. In this case, being the dependent variable a 0-1 dummy, V_{ijt} must be seen as the latent propensity of being in a certain status. For instance, in the case of labour force participation, V_{ijt} can be seen as an inverse monotonic transformation of the worker's valuation of the job he/she has or would have at time t if he/she participated in the market. If this value is above a certain threshold, the individual will participate. Then the probability of participating can be estimated using a fixed-effect logit model, by maximising the conditional likelihood (see *e.g.* Baltagi, 1995).

In contrast to the case of wage and employment security, however, when labour force participation and/or unemployment status are used as dependent variable, decomposing the training stock between training with previous employers and other training would not completely solve the identification problem. As discussed above, one can be guaranteed that the estimate of the effect of training taken with

previous employers is upward biased — because of match-specific selectivity — only if the true effect of training taken with the current employer is positive and greater than the true effect of training with previous employers (that is if $\gamma < \delta$ in [1]). Therefore, strictly speaking, to talk of a positive impact of training, it would be necessary to check that the estimated effect of training taken with previous employers were not only significantly different from zero but also greater than that of training with the current employer. However, due to the fact that the effect of training with the current employer cannot be identified (since, by definition, training with the current employer can be different from zero only if the worker is employed), it cannot be excluded that the estimates are upward biased because of selectivity. This is the reason why the analysis of individual labour market participation and unemployment is complemented with the analyses of the impact of training with previous employers on wage premia and job security where the training stock can be meaningfully decomposed in its relevant components and, at least in the most relevant cases, the estimated impact of training taken with previous employers turns out greater than that of training with the current employer as well as significantly different from zero.

As regards to employment insecurity, which is defined on a 1-6 Likert scale, following the literature on job satisfaction, one could use the same methodology to estimate a fixed-effects linear model (Heywood *et al.*, 2002) or a fixed-effects logit model (Winkelmann and Winkelmann, 1998), by collapsing the measure of job security into a dichotomous variable. However, neither of these methods is ideal, since in the first case the qualitative (or at least double censored) nature of the data is not taken into account, while in the second case a great deal of information is thrown away. An alternative, suggested by Heywood *et al.* (2002), is to estimate [1] in first differences — possibly using observations at relatively distant dates. The use of first differences eliminates individual effects and allows choosing within a wider set of models. In Chapter 4 of OECD (2004), three types of models (linear, Gaussian interval regression and generalised ordered probit models) have been estimated with qualitatively identical results.

A standard ordered probit model can be defined as follows: let D_{it}^* be a latent variable and D_{it} the observable discrete dependent variable which can take N ordered values $k \in (1, 2, \dots, N)$ (with the smallest value normalised to 1 for convenience) depending on the realisation of D_{it}^* and according to the following rule:

$$\left\{ \begin{array}{ll} D_{it} = k, 1 < k < N & \text{if } \sigma_{k-1} < D_{it}^* \leq \sigma_k \\ D_{it} = k = 1 & \text{if } D_{it}^* \leq \sigma_k \\ D_{it} = k = N & \text{if } D_{it}^* > \sigma_{k-1} \end{array} \right.$$

where $(\sigma_1, \dots, \sigma_{N-1})$ are $N-1$ unobservable ordered thresholds, and D_{it}^* is assumed to be normally distributed. In the case of the change in job security, $D_{it} \in (-5, -4, \dots, +5)$ and $D_{it}^* = V_{ijt} - V_{ijt-1}$. However, for any given lagged level of perceived job security only 6 values can be observed. The generalised version used in Chapter 4 of OECD (2004) adjusts the likelihood function to take into account that certain values of D_{it} cannot be observed and that the values that can be observed depends on the lagged value of perceived employment security s_{it-1} :

$$\left\{ \begin{array}{ll} D_{it} = k, 2 - s_{it-1} \leq k \leq 5 - s_{it-1} & \text{if } \sigma_{k-1} < D_{it}^* \leq \sigma_k \\ D_{it} = k = 1 - s_{it-1} & \text{if } D_{it}^* \leq \sigma_k \\ D_{it} = k = 6 - s_{it-1} & \text{if } D_{it}^* > \sigma_{k-1} \end{array} \right.$$

By contrast, in the Gaussian interval regression model used in Chapter 4 of OECD (2004), D_{it} is observed according to the following rule:

$$\left\{ \begin{array}{ll} D_{it} = k, 2 - s_{it-1} \leq k \leq 5 - s_{it-1} & \text{if } k - 1 < D_{it}^* \leq k \\ D_{it} = k = 1 - s_{it-1} & \text{if } D_{it}^* \leq k \\ D_{it} = k = 6 - s_{it-1} & \text{if } D_{it}^* > k - 1 \end{array} \right.$$

with D_{it}^* is assumed to be normally distributed. The difference between the two models is that in the latter D_{it} has a cardinal interpretation and all intervals have length 1 (except those that are bottom and top coded), while in the former D_{it} has only an ordinal interpretation since the thresholds are not known. However, one drawback of these models is that left- or right-censoring of s_{it-1} is in principle ruled out. This implicit assumption is in contrast with the normality of D_{it}^* , since both D_{it} and s_{it-1} are generated by transformations of [1]: assuming normality of ε_{it-1} implies that s_{it-1} is bottom-coded at 1 and top-coded at 6. The normality assumption must be regarded, therefore, only as an approximation. For this reason, it is useful to check the results by estimating also a linear model, which implies no distributional assumptions (see Bassanini, 2004, for the results from this sensitivity analysis).

In the case of the analysis of the effect of training on the re-employment prospects of displaced workers, the natural adaptation of a fixed effect framework would be to study the effect of training on the length of the unemployment spell after dismissal within a Cox's proportional hazard model (see, for instance, Parent, 1999, for an application of this model to job mobility). Unfortunately, due to the lack of cross-country comparative datasets where many spells of unemployment are observed for a large portion of the sample, this model cannot be implemented here. Chart 4.13 is therefore based on a simple econometric specification that can again be derived by first-differentiating [1] over a two-year period. The problem with using this specification is that the sample (being limited to displaced workers) is truncated. In fact, the probability of being displaced depends on the realization of V_{ijt-2} . Obviously, the greater the fixed effect μ_i the lower the greatest value ε_{it-2} that will induce a lay-off. If the random disturbance is independently distributed, then conditional on being dismissed in $t-2$, $(\varepsilon_{it} - \varepsilon_{it-2})$ will be positively correlated with the fixed effect and, consequently, is likely to be correlated with the stock of training taken with the last employer. As a consequence, a sensitivity analysis is presented in Annex 1 to check the importance of this source of selectivity bias. Although the absence of appropriate instruments for training makes it impossible to control for selection on unobservables, matching methods based on the propensity score — Nearest Neighbour matching and Gaussian Kernel matching based, in both cases, on a common support; see Heckman *et al.* (1997) — are used in this analysis to correct for selection on observables, including the characteristics of the last job. In fact, it can be argued that the characteristics of the lost job (including the wage) are sufficient to capture unobserved individual ability, at least insofar as selection into treatment is concerned. Therefore, if these variables are included among the explanatory variables, it can be argued that treatment effects are orthogonal to the probability of treatment conditional on observables. Hence, estimates conditional on the propensity score are unbiased and consistent estimates of the average treatment effect on the treated (the average effect of training on those receiving it).

BIBLIOGRAPHY

- Baltagi, B.H. (1995), *Econometric Analysis of Panel Data*, Wiley, Chichester.
- Bassanini, A. (2004), “Improving Skills for More and Better Jobs? The Quest for Efficient Policies to Promote Adult Education and Training”, paper presented at the European Commission Conference on Quality and Efficiency in Education, Brussels, May.
- Booth, A. and M. Bryan (2002), “Who Pays for General Training? New Evidence for British Men and Women”, *IZA Discussion Paper*, No. 486, April.
- Gerfin, M. (2003), “Firm-sponsored Work Related Training in Frictional Labour Markets: An Empirical Analysis for Switzerland”, *Universität Bern Diskussionschriften*, No. 03-17, October.
- Heckman, J.J., H. Ichimura and P. Todd (1997), “Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program”, *Review of Economic Studies*, Vol. 64, pp. 605-654.
- Heywood, J.S., W.S. Siebert and X. Wei (2002), “Worker Sorting and job Satisfaction: The Case of Union and Government Jobs”, *Industrial and Labor Relations Review*, Vol. 55, pp. 595-609.
- Loewenstein, M.A. and J.R. Spletzer (1998), “Dividing the Costs and Returns to General Training”, *Journal of Labor Economics*, Vol. 16, No 1, pp. 142-171.
- Loewenstein, M.A. and J.R. Spletzer (1999), “General and Specific Training: Evidence and Implications”, *Journal of Human Resources*, Vol. 34, No. 4, pp. 710-733.
- OECD (2004), *Employment Outlook*, Paris.
- Parent, D. (1999), “Wages and Mobility: The Impact of employer–Provided Training”, *Journal of Labor Economics*, Vol. 17, No 2, pp. 298-317.
- Winkelmann, L. and R. Winkelmann (1998), “Why Are the Unemployed So Unhappy? Evidence from Panel Data”, *Economica*, Vol. 65, pp. 1-15