

DISENTANGLING FACTORS BEHIND TRAINING PARTICIPATION IN ITALY*

Giuseppe Croce and Massimiliano Tancioni*

1. Introduction

In recent years, Government institutions assumed education and training of adults as a major leverage to pursue the structural adjustment of the economy as well as to improve the labour market prospects of individuals. In particular, workers' training has been conceived as a remedy to counteract the widening gaps between skilled and unskilled persons.

Nonetheless, further investigations are requested to support this policy strategy and to design proper training measures. Groups facing poor training opportunities and factors affecting training participation have to be carefully detected for a more effective implementation of targeted policies.

The economic analysis should also attempt to distinguish whether inequalities in training participation reflect efficient investments or, on the contrary, imply inefficiencies too, as the rationale for public intervention in the two cases is different (Snower and Booth 1996, Lynch 2003, OECD 2004).

* We acknowledge ISFOL and, in particular, E. Mandrone for providing us with data from the survey Plus (2005). Micro-data use authorization codes Isfol PLUS 2005/0022-0023.

* University of Rome, La Sapienza.

As a matter of fact, the empirical evidence exhibits large differences in participation to training activities among various groups of workers. As such differences exist also in countries where the overall participation rate is higher, they are expected to be even more relevant in countries with low training incidence, as Southern countries of Europe.

Explaining training participation has to be regarded as a tricky task since both workers and employers can play a role in training investment decisions. In other words, the observed pattern of participation derives from joint decisions and it is not easy to distinguish the factors determining the workers willingness to participate to training and the employers propensity to finance it. Oosterbeek (1998) made clear that, because of lack of information, estimates of training participation mostly refer to a reduced form model, whereas a structural model would be requested in order to disentangle factors impinging on workers' and employers' choices.

Employing the information provided by a new survey conducted on a large sample of individuals, this paper proposes an analysis which is potentially capable of identifying the structure of the training choice, thus of estimating a model of the training choice in its structural form. As long as our approach approximates such a structure, policy implications can be derived from the results of the analysis.

In particular, we are interested to evaluating the decision criteria the employers adopt to select eligible workers for a training programme sponsored by the firms. It is a well-known fact that they normally play a prominent role in promoting training activities (Bassanini et al. 2005). Then, a relevant question is whether firms selectivity, which is assumed to depend on differences in their private return on training different groups of workers, also reflects the social return or, conversely, it deviates from it. In the latter case public interventions aimed at favouring training opportunities of disadvantaged groups could bring about some reduction of inequality together with efficiency improvements, whereas in the former the standard trade-off between equality and efficiency would arise.

A general principle in training policy maintains that worker and employer have to sustain the largest part of training costs as they reap most of its benefits. Nevertheless, a number of market failures can justify public interventions. From a general point of view public policy is aimed at increasing training investments at large even if in most cases the public intervention would resolve more efficient if implemented as a set of measures targeted to selected groups,

that is those groups of workers who enjoy poor training chances. This implies that a deeper empirical analysis is required in order to establish if underinvestment primarily depends on worker's and/or employer's attitudes. Furthermore, a careful assessment should reveal if low training implies inefficiencies besides inequality problems.

The most relevant feature of our dataset is that it provides information not only on training participation but also on its financing. This information must be considered cautiously as individuals could not perfectly perceive which subject (employer, government and other public agencies, individual themselves) actually contribute to sustain the training direct and indirect costs and how large is their respective cost share. Moreover, the items included in the questionnaire to specify the source of financing only permit an approximate answer. However, in our analysis we do not rely on punctual information on financing as at this stage we merely need to distinguish between the training provided by the employer and that acquired by the worker from other sources.

Accordingly, we group cases in two categories: "internal training", corresponding to training organised and/or financed by the employer, and "external training", i.e the training financed by local and regional governments through vouchers, by the European Social Fund, by the worker himself or free for other reasons.

Furthermore, we also exploit additional information concerning workers who did not participate to training activities but declare to have applied for a course. These workers can be considered as "rationed" workers as they searched for training but their demand didn't match any suitable offer¹.

2. A structural model of training participation

Empirical evidence across countries reveals that employers play a crucial role in financing and providing training opportunities to their employees. Internal training always requires a joint decision by the employer and the worker. Available information usually reports only whether training occurred or not, without any further information allowing to

¹ Oosterbeek (1998) and Leuven and Oosterbeek (1999) exploit analogous information regarding workers who "wanted" to receive training but did not do so.

distinguish between the worker and the employer's behaviour. Based on this information, at best only reduced form models of training participation can be estimated (see for example Arulampalam et al. 2003). Even if factors associated to low (high) participation can be detected, it is not possible to establish if they should be attributed to the workers or to the employers choices or to both.

Few recent papers tried to overcome this limitation and to estimate structural models of participation. Oosterbeek (1998) firstly proposed to identify training "demand" by workers and "supply" by firms by exploiting the fact that IALS asks respondents who did not participate to any training if they would like/wanted to do it (the "rationed" workers). Leuven and Oosterbeek (1999), Bassanini and Ok (2004) and OECD (2003) provide further applications of this scheme. All these papers are based on data from IALS for the '90s.

In OECD (2003) it is assumed that firms acquire training in an upstream market and, correspondingly, resell it to the workers in a downstream market. Then, at this second stage, firms supply training while workers demand it. In such context participation to training as well as rationing represent training demand. Nevertheless, participation to internal training has to be attributed also to training supply by firms. However, such a scheme represents a partial representation of the training opportunities as it neglects that workers can acquire it also through channels that are external to the firm, as with public agencies or private providers.

The exclusion of external training would be an arbitrary limitation which hampers a satisfactory identification of the demand for training by the workers. For this reason we adopt a scheme (see OECD 2003 and Bassanini and Ok 2004) where external training, besides internal, is explicitly considered. Furthermore, we use a different and more recent dataset provided for by *Plus*, a survey conducted by ISFOL in 2005. It represents a new dataset which allows us to apply such an analysis to Italy for the first time.

Our starting point is that participation decisions by employers and employees depend on their private net return on it. Accordingly, we can specify the following two equations

$$\begin{aligned} y_f &= \alpha_f + \beta_f \mathbf{x} + \varepsilon_f \\ y_w &= \alpha_w + \beta_w \mathbf{x} + \varepsilon_w \end{aligned} \quad (1)$$

where y_f and y_w represent, respectively, the firm and the worker's net return on training the individual worker (subscript i has been omitted), \mathbf{x} is the vector of the variables measuring observed characteristics of workers, jobs and firms, $\boldsymbol{\beta}_f$ and $\boldsymbol{\beta}_w$ are the vectors of coefficients, α_f and α_w are group-specific constant terms and ε_f and ε_w are the error terms.

We assume that the firm will offer training opportunities to the worker if and only if it benefits from it, that is $y_f > 0$. On the other hand, the worker will demand for training (that is he will accept training offered by the employer or access training activities offered from outside the firm) if and only if he finds it convenient, which is the case when $y_w > 0$. Apart from this case, he will never accept to train himself neither inside the firm nor outside.

Four different situations can be distinguished by jointly considering the net gains of the firm and the worker:

$$\begin{aligned}
 (i) \quad & y_f > 0 \quad \text{and} \quad y_w > 0 \\
 (ii) \quad & y_f \leq 0 \quad \text{and} \quad y_w > 0 \\
 (iii) \quad & y_f > 0 \quad \text{and} \quad y_w \leq 0 \\
 (iv) \quad & y_f \leq 0 \quad \text{and} \quad y_w \leq 0
 \end{aligned} \tag{2}$$

Though exact measures of returns y_f and y_w are not usually reported in any dataset, we are able to infer from sample selection whether the worker and employer's net returns are positive or negative. This, in turn, makes it possible to identify which is the underlying situation corresponding to each observation concerning training participation. Following Oosterbeek (1998) we assume that the employer is not able to impose training to the worker. This also follows from the hypothesis that participation strictly reflects the existence of a positive return. Participation to internal training always requires that both conditions under (i) in (2) are satisfied.

As we lack information on cost borne by the worker and on the specific training content, differences in net return on internal and external training cannot be estimated. Then we assume that each worker, characterised by a specific set of values of the relevant variables,

can gain the same net benefit from the internal and the external training. This hypothesis, even if unavoidably strong, is implicit also in previous works and corresponds to assuming that total demand for training by each group of workers - which is negatively related to the (implicit) price - can be represented by a unique curve. Participation to free or subsidised training has to be included in the demand for training as it ever implies some costs in terms of time and effort, although these elements are seldom recorded by statistical information. On the other hand, supply of training is assumed to increase with the (implicit) price. Finally, both demand and supply slopes are assumed to be invariant across groups of workers (for details see OECD 2003 and Bassanini and Ok 2004).

On the basis of these assumptions, we can attribute each observation on training participation to one of cases (i)-(iv) in (2). When internal training is observed, it can be inferred that situation (i) occurred. We label this case as *internal equilibrium* as it corresponds to the situation where both the firm and the worker find training profitable. On the other hand, external training as well as rationing correspond to situation (ii). In such a case, the worker would receive training but no opportunity is offered to him by the firm so that he has to resort to external providers. Finally, situations (iii) and (iv) apply when no participation, neither inside nor outside the firm, is observed, as in both cases training does not give rise to any positive gain for the worker. Because of data limitation it is impossible to distinguish case (iii) from case (iv). However, it is worth noticing that in case (iii), unlike case (iv), employers do offer training but this does not take place because of worker's reluctance.

Then we can define the dichotomous variables z_f and z_w , where z_f takes value 1 if the net return y_f is positive and zero otherwise, as well as z_w takes value 1 when y_w is positive and zero otherwise.

Contrary to the usual estimates of training probit models based on information on participation only, we are able to carry out a more structured analysis aimed at separating the effects of the explanatory variables on the worker's and firm's decisions. Given the scheme adopted here, this corresponds to the identification of demand and supply of training.

We define two probit equations. In the first one, the dependent variable equals 1 if the worker underwent training during the three years before the interview, either inside or

outside the firm, or she/he declares himself to be rationed – situations (i) and (ii) – and zero otherwise – situations (iii) and (iv) – . According to our scheme, this equation should capture the effects of each individual characteristic on the probability that training occurs or that worker reports some rationing. This corresponds to estimate the vector of parameters for the individual characteristics which defines:

$$P(\text{internal or external training or rationing occur}) \equiv P(z_w = 1). \quad (3)$$

In other words, from this probit equation we get an estimate of the vector of coefficients β_w measuring how factors affect worker's willingness to take training, that is her/his *demand* for training.

On the other hand, the second probit equation, which applies only to the sub-sample of trained and rationed workers (those with value 1 in the first equation) provides an estimate of the effects of the independent variables on *internal equilibrium*. In this case the dependent variable takes value 1 when internal training occurred – situation (i) – and zero in case of external training or rationing – situation (ii). This corresponds to the estimate of the coefficients on the individual characteristics affecting the probability of internal equilibrium:

$$P(\text{internal training occurs given the worker's demand}) \equiv \Pr(z_f = 1 | z_w = 1). \quad (4)$$

Though we are not able to directly estimate the coefficients vector β_f , representing the role played by individual characteristics on the employers' willingness to train, their sign can be inferred by comparing the coefficients derived from the two probit equations. By this way we are able to disentangle demand and supply effects shaping the pattern of training participation. This represents a valuable step forward in explaining the distribution of training across different groups of workers.

3. Empirical model and estimation strategy

Operationally, the basic empirical formulation of our model is the bivariate probit model:

$$\begin{aligned} y_{if} &= \alpha_f + \boldsymbol{\beta}_f \mathbf{x}_{if} + \varepsilon_f, & z_{if} &= 1 \text{ if } y_f > 0, & z_{if} &= 0 \text{ otherwise} \\ y_{iw} &= \alpha_w + \boldsymbol{\beta}_w \mathbf{x}_{iw} + \varepsilon_w, & z_{iw} &= 1 \text{ if } y_w > 0, & z_{iw} &= 0 \text{ otherwise} \end{aligned} \quad (5)$$

where $[\varepsilon_{if}, \varepsilon_{iw}] \sim \text{BVN}(\text{bivariate normal})[0,0,1,1,\rho]$. Notice that the standard univariate case arises if $\rho = 0$, whose occurrence is testable employing the Lagrange multiplier statistic on $H_0 : \rho = 0$ ². Given the approach employed here, we do not expect to find independence between the two equations, as they are estimated employing (partially) overlapping sample information.

Differently from standard structural models, instead of imposing theory-based coefficient restrictions, identification is obtained from sample selection³. In other terms, in the initial estimate we do not restrict neither the variables nor the signs of the coefficients of the two equations. This is possible given our theoretical apparatus briefly sketched in the preceding section, which implies that identification can be obtained by discriminating the possible dichotomous outcomes on z_i .

In order to highlight the differences between our structural approach and the standard reduced-form models estimates, we start our analysis by estimating a standard univariate probit model in which the dependent variable is 1 if training occurs and 0 otherwise. Results are thus compared with those from the bivariate probit model (5), estimated on the same set of regressors.

Firstly, we estimate a bivariate probit with a very general specification of the regressors space; then we get the final estimate by implementing a reduction process based on statistical information only⁴. After reduction, the final specification of our model may result structural also in the standard meaning, as it can show a different parameters structure for the two

² For the details of the test see Greene (2000).

³ The estimator is proposed by Wynand and van Praag (1981). For an extensive application which uses sample selection see Boyes, Hoffman, and Low (1989).

⁴ This specification search approach is generally referred to as a General to Specific reduction process (GetS, Hendry, 1995).

equations. It is important to highlight that this result is not theory-driven, as theory affects the identification stage for sample selection only. In such a case – and assuming cross-dependence in the error structure – the appropriate estimator is the Seemingly Unrelated Regression (SUR) bivariate probit.

4. Sample selection and the definition of the independent variables set

The Isfol *PLUS* survey contains information on the characteristics of 40386 individuals, selected according to their status of participation to the labour market (active unemployed, employed, pensioners). The employed group consists of 16397 individuals, of which 12736 (nearly 60%) are dependent workers. Given our aim of identifying training supply and demand, we restrict our attention to the latter subset only. The high variability and idiosyncrasies emerging for the younger in the Italian labour market suggests of selecting individuals aged 20 or more only, which leads to a further sample reduction (12446 dependent workers). Moreover, since in the survey questionnaire the individuals are asked to answer on the basis of a three years training participation record, we further restrict our sample to those declaring an employment status persisting for three years or more. This guarantees that the sample, other things equal, is balanced in terms of training opportunities of the representative worker. Given the last restriction, the sample is composed of 12050 individuals, which we define “operational”.

After having imposed our sample selection strategy discussed in section 2 to the operational sample, we end up with the following data structure:

- i) 3205 individuals (26.3% of the operational sample) participating to an internal training programme (case *i* in 2);
- ii) 5939 individuals (49.3% of the operational sample) participating to an internal/external training programme;
- iii) 6130 individuals (50.9% of the operational sample) being trained (internally or externally) or not being trained even having declared their availability to participate to a training programme. This group defines case *ii* in (2);
- iv) 191 (1,6%) are those who have not participated to a training programme, irrespective of the internal/external distinction. This implies that our identification

based on sample selection strongly relies on the internal/external distinction, which is assumed to render a demand-rationing meaning.

Concerning the definition of the independent variables set, for our starting estimate we select a very general set, in which individual, job-specific and firm's characteristics are considered. Our starting set consists of twelve variables, of which three are continuous and nine dichotomous.

The continuous variables are:

- 1) age of the employee (*age*);
- 2) number of years of work with the present-time employer (*nyll*);
- 3) size of the firm in which the individual works, defined by the number of dependent workers in the firm (*f_size*).

The dichotomous variables are:

- 1) sex of the employee (*f*, being *m* the control variable);
- 2) the employee is the head of the family (*head*);
- 3) presence of family members economically depending from the employee (*members*);
- 4) regional area of residence of the employee (*nw*, *ne*, *south*, being *center* the control variable);
- 5) level of study of the employee (*study1*, *study3*, *study4*, *study5*, being *study2* the control variable);
- 6) economic sector to which the firm belongs to. We consider 12 sectors: agriculture (*agric*), industry (*ind*), public utilities (*publ_ut*), constructions (*constr*), trade (*trade*), transports and commerce (*tr_comm*), financial (*fin*), government (*gov*), educational (*edu*), health (*health*), other services (*oth_serv*);
- 7) duration of the job contract, temporary or permanent (*temp_c*, being *perm_c* the control variable);
- 8) part-time worker (*p_time*, being *f_time* the control variable);
- 9) job level, defined in five specialisation classes from high to low (*job_h*, *job_mh*, *job_ml*, *job_l*, being *job_m* the control variable).

Since the continuous variables are entered both linearly and squared in order to take into account possible nonlinearities among the dependent variable and the specific regressor, the actual number of continuous variables is six. Given the levels of aggregation considered for the dichotomous variables and considering those omitted for normalisation, the actual number of dichotomous variables is 28.

The total number of independent variables, once the constant term has been introduced in the explanatory variables space, is thus 35. Table 1 gives a means-based sample description for the set of regressors employed in the starting specification, distinguishing between “demand” and “internal equilibrium” and the respective dichotomous outcomes (0 and 1).

5. Estimation results

Estimation results are summarized in four tables, two for the starting specification (unrestricted explanatory variables space) and two for the final specification (restricted after GetS reduction).

Each estimation result table is organised as follows: The first column contains the estimates for the univariate probit and the relative z -statistics. The second and third columns contain the estimates for the bivariate probit model. Below the tables the number of observations, the log-Likelihood value and the LR test results for the hypothesis of off-diagonal zero error correlation are reported. For the final specification, its statistical viability is summarized by a standard LR test for exclusion restrictions. Being the restrictions derived by a structured reduction process driven by statistical information only, they are always accepted.

The LR test for zero off-diagonal correlation rejects the null hypothesis in both the unrestricted and the restricted formulations, indicating that the bivariate probit is the appropriate model.

The results of the univariate model (Tables 2 and 3, first column) illustrate the effects of regressors on the probability that training occurs. According to them, training is a slightly less frequent event for women than for men (Table 2). Highly educated workers face a higher probability of training. Participation increases with firm size and, contrary to what is

expected, also with age. Moreover, it substantially decreases in case of part-time employment while it is affected in a lesser degree in case of temporary employment.

The estimate of the bivariate model (Tables 2 and 3, second and third columns) makes clearer the causal relationships underlying such results. Indeed, it allows us to distinguish whether the effects of regressors have to be attributed to demand or supply, or both. In this regard, for example, we find that the disadvantage of women and temporary employees mainly depend on employers' behaviour.

In our approach the estimates of the effects on training supply can only be derived by comparing the effects of individual characteristics on demand with those on internal equilibrium. When the effect on demand is positive (negative) and that on internal equilibrium is negative (positive) or absent, it can be concluded that supply had a negative (positive) shift (see OECD 2003 and Bassanini and Ok 2003). When the effects on demand and internal equilibrium have the same sign, it can be more difficult to establish the sign of the impact on supply (in other terms, demand and supply are weakly identified) and it will remain uncertain unless differences in coefficient dimensions for demand and equilibrium are large enough in module. In this case we can roughly infer which is the direction of the effect on training supply. In sum, with due cautions, we can exploit our results to draw new explanations of the observed training pattern.

The results we obtained from the estimate of the final specification (Table 3) suggest that **females** do not demand less training than males. However, the negative change in the internal equilibrium means that women suffer from some rationing inside the firm. Then, low participation of women mainly arises because of employers' reluctance to train them. In its turn, this can depend on higher turnover or on discrimination. This finding confirms previous studies reporting that females demand is greater or at least similar to that of their male peers but it is constrained by a shortage of training supply⁵.

As expected, training demand steeply increases with the worker's **educational level** while, more surprisingly, no similar effects of education on the supply side are noticeable.

⁵ Comparable outcomes are reported by Oosterbeek and Leuven (1999), who estimated a tobit model with censoring, with the dependent variable representing the quantity of training, in their study conducted on IALS data for Canada, Netherlands, Switzerland and United States in the mid-90's; by OECD (2003), based on the same dataset related to a larger number of advanced countries and by Bassanini et al. (2005) on ECHP data on European countries for the period 1995-2001. On the contrary, Arulampalam et al. (2003), estimate for Italy a greater probability of training for women.

Highly educated workers, who can reap the largest benefits from training, do not find adequate opportunities in their firms. Training supply by the employers seems to decrease for the most educated while stays constant or becomes more abundant for less educated employees. Therefore complementarity between possessed education and training, which represents a widely accepted fact (Brunello 2001, Arulampalam et al. 2003), results to be relevant on the demand side while is not confirmed as far as employers preferences are considered.

Explanations of this fact can rely both on the benefits and costs elements. On the one hand, one can argue that occupational needs of Italian enterprises are concentrated on low and intermediate positions, because of the relative scarcity of innovative activities in the economy. However, explanations cannot rely only on national-specific factors as similar results are obtained also for other countries (see Oosterbeek and Leuven 1999, OECD 2003 and Bassanini et al. 2005). On the other hand, it can also be presumed that firms cannot afford to provide inside training for high skilled as this would imply sophisticated and costly requirements. For this reason external training tends to substitute the internal one, and the role of employers becomes less prominent.

In short, these findings suggest that the low educated do not suffer from a shortage of training chances due to employers' selectivity. On the contrary, low participation depends on workers' weaker preference for it. Then, training measures should be addressed to workers rather than to firms. Nevertheless, other policies, like adults education and active labour policies or monetary transfers could be more effective substitutes for training policy to help people with low education.

Both demand and internal equilibrium rise with respect to age, although the coefficients of squared **age**, both negative and significant, reveal a non-linearity. Comparison of coefficients of the demand and internal equilibrium equations reveals that also the effect of age on training supply must be positive. This finding is not consistent with standard human capital theory which predicts that older individuals are less likely to take training. It can be argued that the higher turnover experienced by young workers discourage employers from offering training chances to them. At the same time, also workers tend to postpone investments given initial employment instability.

Evidence from earlier studies appears somewhat mixed. In Bassanini et al. (2005) the age-training profile results to be downward-sloped. Oosterbeek and Leuven (1999) find a negative and significant effect of age on the workers' demand. OECD (2003), on the other hand, shows that an increasing effect of age on employer's offer of training is present, and Arulampalam et al. (2003) find that Italy is the only country where age does not affect training probability.

Demand increases with **current tenure**, that is the number of years of work with the present-time employer (*nyll*). This effect parallels and strengthens that of age. On the contrary, it is more difficult to ascertain its effect on the supply. At the initial stage of the employment relationship uncertainty about the quality and the duration of the matching makes the workers less eager to invest in skills acquirement. Afterwards, their investment propensity increases as the relationship proves satisfactory and the employment prospects become less volatile.

Also the **employment contract** affects training investments. Our analysis confirms that, on the one hand, employees without a permanent contract demand as much training as their permanent colleagues but, on the other hand, they are short of training chances inside the firm. On the contrary, according to the estimates of univariate models (see Table 3, first column and also Arulampalam et al., 2003), training probability for Italian workers seems to be unaffected by the duration of the contract. However, such finding is spurious, as estimates of univariate models fails to separate the effects on demand and on supply. The bivariate model reveals that temporary workers do not enjoy enough employer-sponsored training even if this could be beneficial to them. Firms choices, in this case, are negatively affected by poor prospects of recuperating the training cost, due to the shorter expected duration of employment. For this reason the socially efficient result can be far from being attained. Temporary employment seems to imply not only inequality in training participation but also a loss of efficiency. On this respect policy implications can be drawn along two main lines. First, policy measures could be addressed to favour training of temporary workers outside the firm. Second, labour policies should make easier the transition from temporary to more stable employment.

Part-time workers appear to be in a different situation as they exhibit a lower demand respect to those working full-time while it is uncertain the effect on the supply side (this

result is very close to that provided by Bassanini et al. 2005). Lower demand likely depends on the same factors preventing these employees from working full time. However, this hypothesis should be further verified by distinguishing between voluntary and involuntary part-timers. Indeed, OECD (2003) reports that involuntary part-timers prefer training as much as workers with full-time contract do.

As far as **regional areas** are concerned, Northeast and Northwest – which are confronted to the other regional areas in our final specification – display an opposite pattern, as in the former area participation is driven by stronger demand by workers whereas in the latter one a higher supply emerges. Tentative explanations can point to differences between regional labour markets relatively to structural characteristics as labour mobility and wage compression. More intense mobility and a less compressed wage structure in the Northeast likely shift the incentive to invest in training from the employers to the workers. The reverse could occur in the Northwest. Besides this, differences in firms' technological and organisational characteristics as well as in managerial culture, not fully captured by firm size and industrial dummies, can also contribute to explain this outcome.

Job characteristics influence workers' willingness to take training. From the first probit we derive that, as expected, demand increases with the rank of the job. High level (managers, professionals and highly specialised technicians) and medium-high level (teachers and other technicians) workers take training more frequently than medium level (the reference group, comprehending clerks and specialised workmen) ones. On the other hand, medium-low level (call center operators, service workers, shop assistants, craftsmen, plant and machine operators and generic workmen) and low level (elementary occupations) workers are less involved in training activities. Firms do not follow a similar selection criterion. Supply of training is lesser for higher levels while no clear effect can be inferred for the lower levels⁶. Then, this finding parallels the effect of education discussed above.

Higher hierarchical positions in large and medium enterprises require sophisticated (and mostly generic) knowledge, which workers more often acquire by themselves. On the other hand, higher positions in small firms are mainly characterised by tacit knowledge, which is accumulated by experience and informal relationships rather than through formal courses. In both cases internal training does not play a primary role.

⁶ OECD (2003) estimates based on IALS find the reverse outcome.

Training probabilities are affected also by the **firm size**. Both its effect on the worker demand and on internal equilibrium is positive and significant. However, the inclusion of the squared size reveals some non-linearity in the size-training profile.

Finally, the employees in the **industries** included in our final specification (transport and communications, finance, government, education, health, other services) demand training more frequently than their peers in other industries. The coefficients of education and other services in the internal equilibrium equation are not significant, meaning that the shift of supply could be negative in such cases. On the contrary, employment in the financial sector gives rise to the highest effect.

6. Conclusion

The observed pattern of participation to training derives from joint decisions by the workers and the employers and it is not easy to distinguish the factors determining the workers willingness to receive training and the employers propensity to finance it. Because of lack of information, estimates of training participation usually refer to a reduced form model, whereas a structural model would be requested in order to disentangle factors impinging on workers' and employers' choices. Employing the information provided by a new survey conducted on a large sample of individuals, the paper provided an estimate of a model of the training choice in its structural form. This represents a valuable step forward in explaining the distribution of training across different groups of workers.

Differently from standard structural models, in the initial estimate we did not restrict neither the variables nor the signs of the coefficients of the two equations. Firstly, we estimate a bivariate probit with a very general specification of the regressors space; then we get the final estimate by implementing a reduction process based on statistical information only. In order to highlight the differences between our structural approach and the standard reduced-form models estimates, we estimated both a standard univariate and a bivariate probit model on the same set of regressors.

Our findings suggest that employers are reluctant to train women and temporary workers, though they would like to receive as much training as their peers would do. Highly educated

workers, whose training can yield the largest benefits, are prone to acquire it outside the firm as they do not find adequate opportunities inside. At the same time, the low level of training participation of the less educated depends on workers' weaker preference rather than on employers' selectivity. Contrary to the prediction of the human capital theory, the age-training profile is proved to be upward-sloped. Indeed, both the demand and the supply of training increase with age. Part-time workers exhibit a lower demand respect to those working full-time.

Further substantial developments of this line of research would require more information on training costs, e.g. through integration of different sources, as surveys on workers and employers and administrative data.

References

- Arulampalam W., Booth A.L., Bryan M.L. (2003), *Training in Europe*, IZA discussion paper No.933.
- Bassanini A., Booth A., Brunello G., De Paola M., Leuven E. (2005), *Workplace training in Europe*, Report to the Conference of the Fondazione De Benedetti, Venice.
- Bassanini A., Ok W. (2004), *How do individuals' incentive to invest in human capital vary across groups?*, Joint EC-OECD Seminar on Human Capital and Labour Market Performance, Brussels.
- Brunello G. (2001), *On the complementarity between education and training in Europe*, IZA discussion paper No.309.
- Greene, (2000), *Econometric Analysis*, Prentice Hall International.
- Hendry, (1995), *Dynamic Econometrics*, Oxford University Press, Oxford.
- Leuven E., Oosterbeek H. (1999), *The demand and supply of work-related training: evidence from four countries*, *Research in Labor Economics*, 18, 303-330.
- Lynch L.M. (2003), *Comment*, in J.J. Heckman, A.B. Krueger, *Inequality in America. What Role for Human Capital Policies?*, MIT Press, Cambridge, Massachusetts.
- OECD (2003), *Employment Outlook*, Paris.
- OECD (2004), *Employment Outlook*, Paris.
- Oosterbeek H. (1998), *Unravelling supply and demand factors in work-related training*, *Oxford Economic Papers*, 50, 266-283.
- Snowder D.J., Booth A.L. (1996), *Government policy to promote the acquisition of skills*, in Booth and Snowder (eds), *Acquiring skills*, Cambridge University Press, Cambridge.

Table 1: Descriptive statistics (sample means) for the set of regressors included in the starting specification

Variable	demand		Internal equilibrium	
	Dep=0	Dep=1	Dep=0	Dep=1
AGE	39,719	42,328	40,142	43,542
F	0,518	0,539	0,542	0,491
HEAD	0,523	0,489	0,529	0,442
MEMBERS	3,198	3,176	3,196	3,162
NW	0,266	0,231	0,251	0,239
NE	0,217	0,235	0,220	0,244
SOUTH	0,312	0,343	0,327	0,329
STUDY1	0,050	0,005	0,034	0,006
STUDY3	0,556	0,545	0,542	0,575
STUDY4	0,118	0,344	0,205	0,310
STUDY5	0,003	0,027	0,012	0,022
F_SIZE	130,903	144,895	123,897	176,999
AGRIC	0,025	0,008	0,020	0,007
IND	0,225	0,088	0,174	0,105
PUB_UT	0,022	0,017	0,019	0,020
CONSTR	0,043	0,015	0,033	0,016
TRADE	0,162	0,081	0,137	0,076
TOUR	0,045	0,019	0,039	0,012
TR_COMM	0,064	0,059	0,058	0,071
FIN	0,030	0,053	0,029	0,076
GOV	0,114	0,226	0,142	0,251
EDU	0,089	0,222	0,150	0,173
HEALTH	0,048	0,135	0,081	0,123
OTH_SERV	0,118	0,070	0,105	0,063
TEMP_C	0,144	0,115	0,150	0,072
P_TIME	0,202	0,102	0,177	0,078
JOB_H	0,041	0,130	0,075	0,118
JOB_MH	0,119	0,323	0,199	0,288
JOB_ML	0,276	0,079	0,210	0,080
JOB_L	0,069	0,017	0,053	0,014
NYLL	12,931	16,046	13,456	17,441

Table 2: Estimation results from the univariate and bivariate probit models (starting specification)

Regressor	Univariate probit		Bivariate probit			
	Coefficient	z-stat	demand		internal equilibrium	
			Coefficient	z-stat	Coefficient	z-stat
age	0,031	3,01	0,029	2,88	0,041	3,68
age_2	-0,001	-4,32	0,000	-4,11	-0,001	-4,16
f	-0,053	-1,57	-0,029	-0,89	-0,075	-2,13
head	-0,036	-1,08	-0,016	-0,49	-0,019	-0,56
members	0,009	0,73	0,011	0,99	0,008	0,69
nw	0,021	0,56	0,034	0,92	0,076	1,94
ne	0,201	5,19	0,189	5,05	0,174	4,43
south	0,009	0,26	-0,012	-0,33	0,007	0,19
study1	-0,370	-3,32	-0,371	-3,45	-0,266	-2,25
study3	0,475	12,25	0,490	12,94	0,364	8,66
study4	0,810	16,44	0,788	16,2	0,422	8,21
study5	1,287	8,91	1,302	8,73	0,462	4,12
f_size	0,000	3,97	0,000	4,34	0,000	4,62
f_size_2	0,000	-2,71	0,000	-2,82	0,000	-3,36
agric	-0,030	-0,19	0,038	0,25	-0,090	-0,51
ind	0,007	0,06	0,017	0,15	0,106	0,82
pub_ut	0,173	1,23	0,227	1,64	0,229	1,5
constr	0,042	0,31	0,050	0,38	0,047	0,31
trade	0,129	1,1	0,127	1,1	0,102	0,79
tour	0,087	0,66	0,109	0,84	-0,099	-0,65
tr_comm	0,349	2,88	0,409	3,43	0,386	2,9
fin	0,526	4,18	0,571	4,61	0,660	4,84
gov	0,726	6,3	0,764	6,72	0,519	4,07
edu	0,548	4,63	0,573	4,89	0,164	1,25
health	0,830	6,95	0,836	7,08	0,392	2,99
oth_serv	0,208	1,76	0,231	1,98	0,130	0,99
temp_c	-0,034	-0,81	0,002	0,04	-0,145	-3,04
p_time	-0,276	-6,98	-0,289	-7,49	-0,280	-6,29
job_h	0,256	4,6	0,252	4,54	-0,063	-1,16
job_mh	0,326	8,02	0,329	8,09	0,146	3,61
job_ml	-0,418	-10,06	-0,399	-9,82	-0,376	-8,25
job_l	-0,350	-4,69	-0,339	-4,71	-0,339	-3,99
nyll	0,032	5,73	0,032	6,05	0,037	6,38
nyll_2	-0,001	-3,6	-0,001	-3,98	-0,001	-4,68
constant	-1,691	-7,34	-1,622	-7,22	-2,421	-9,69

No observations: 12050

Log-Likelihood: Univariate: -6728.08; Bivariate: -10736.26

LR test of rho=0:

P=(0.000)

Table 3: Estimation results from the univariate and bivariate probit models (final specification, after reduction)

Regressor	Univariate probit		SUR Bivariate probit			
	Coefficient	z-stat	demand		internal equilibrium	
LR test for exclusion restrictions	Chi-sq.(12): 16.21 (P: 0.18)		Chi-sq.(24): 33.17 (P: 0.09)			
age	0,029	2,93	0,028	2,87	0,042	3,9
age_2	-0,001	-4,19	0,000	-4,09	-0,001	-4,3
f	-	-	-	-	-0,062	-2,61
nw	-	-	-	-	0,047	1,88
ne	0,184	6,1	0,175	6,01	0,160	5,07
study1	-0,386	-3,48	-0,381	-3,56	-0,282	-2,45
study3	0,467	12,13	0,482	12,82	0,371	8,88
study4	0,793	16,26	0,776	16,28	0,424	8,65
study5	1,268	8,79	1,284	8,66	0,455	4,2
f_size	0,000	3,91	0,000	4,14	0,000	4,81
f_size_2	0,000	-2,64	0,000	-2,6	0,000	-3,46
tr_comm	0,290	5,51	0,330	6,45	0,285	5,42
fin	0,449	7,06	0,483	7,82	0,552	9,12
gov	0,655	16,7	0,673	17,94	0,403	11,45
edu	0,461	9,17	0,456	10,29	-	-
health	0,750	14,53	0,733	14,67	0,256	5,76
oth_serv	0,136	2,9	0,130	3,49	-	-
temp_c	-	-	-	-	-0,152	-3,7
p_time	-0,283	-7,31	-0,286	-7,6	-0,290	-6,61
job_h	0,248	4,49	0,277	5,74	-	-
job_mh	0,318	7,91	0,341	8,8	0,185	5,56
job_ml	-0,428	-10,88	-0,414	-10,82	-0,369	-8,49
job_l	-0,363	-4,93	-0,340	-4,78	-0,341	-4,03
nyll	0,033	6,12	0,033	6,3	0,037	6,44
nyll_2	-0,001	-3,83	-0,001	-4,11	-0,001	-4,66
constant	-1,639	-8,65	-1,504	-8,15	-2,347	-11,31

No observations: 12050

Log-Likelihood: Univariate= -6736.18; Bivariate= -10752.84

LR test of rho=0: P=(0.000)