An Empirical Analysis of the Decision to Train Apprentices

Samuel Muehlemann — Juerg Schweri — Rainer Winkelmann — Stefan C. Wolter

Abstract. It is a widely held belief that apprenticeship training represents a net investment for training firms, the cost of which needs to be recouped after the training period. A new firm-level data set for Switzerland reveals large variation in net costs across firms and, remarkably, negative net costs for 60 per cent of all firms. We use these data to estimate the effect of net costs on the number of apprentices hired by a firm. The results show that the costs have a significant impact on the training decision but no significant influence on the number of apprentices, once the firm has decided to train. For policy purposes, these results indicate that subsidies for firms that already train apprentices would not boost the number of available training places.

1. Introduction

Every year, almost 70,000 or more than half of the Swiss youngsters who complete their compulsory schooling choose to embark on what is called the dual education system, i.e. a training programme combining vocational education at school with training in and work for a company. In order to guarantee each new cohort of school leavers that there is a sufficient number of firms willing to

---

Samuel Muehlemann — Juerg Schweri — Rainer Winkelmann — Stefan C. Wolter (author for correspondence), University of Berne, Department of Economics, Centre for Research in Economics of Education, P.O. Box 8573, CH-3001 Bern, Switzerland. E-mail: stefan.wolter@vwi.unibe.ch.

The study is based on two surveys financed by the Commission for Technology and Innovation (CTI credit 4289.1 BFS and 5630.1 BFS) with the help of the Swiss Federal Statistical Office and carried out in tandem with a similar survey in Germany conducted by the Federal Institute of Vocational Training (Bundesinstitut für Berufsbildung) in Bonn. We are grateful to an anonymous referee and the editor for valuable comments. The usual disclaimer holds.

LABOUR 21 (3) 419–441 (2007) JEL J24, C25

offer training opportunities, it is important to know which factors affect the number of apprentices hired by a firm. From an economic point of view, it seems obvious that a firm would hire the more apprentices the higher the net benefit of apprenticeship training amounts to. So far, most studies suggested that apprenticeship training results in net cost during the apprenticeship (e.g. Beicht *et al.*, 2004). This would imply that companies need to be able to recoup these net costs by employing graduated apprentices as skilled workers and paying them a wage lower than the value of their marginal productivity, because otherwise they would not offer an apprenticeship programme.

The most recent study of the Swiss apprenticeship system shows, unlike other studies, that about two-thirds of the training firms find it profitable to train apprentices (Schweri *et al.*, 2003). This suggests that for the majority of training firms, the possibility to recoup the investment during the apprenticeship training period is an important factor explaining their availability for training. That not every single company is able to offer apprenticeship posts under the same favourable conditions was subsequently pointed out by Wolter *et al.* (2006), who showed that the potential net costs of non-training firms would be positive and considerably higher than those of training companies, if they were to engage in training apprentices. These observations lead to the hypothesis that, at least in Switzerland where the high flexibility of the labour market puts some restrictions on the possibilities for firms to recoup training expenses after the training period has ended, the net cost of training during the training period is a decisive argument in favour of or against the training of apprentices.

This paper extends the analyses made so far in testing directly whether the (potential) net costs of training during the training period explain the training probability of firms. Furthermore, it is estimated whether the net costs of training also affect the number of apprentices, once the decision to train has been made. We use different types of count data models to account for the specific structure underlying the firms' decision process.

The paper fills a gap in the training literature, as all previous studies (see Section 3) estimating the number of apprenticeship places offered by a firm were not able to include data on the cost and benefit of the apprenticeship training in their analyses because of the lack of these data. The only exception known to the authors is a study by Niederalt *et al.* (2001), using, however, a very small and non-representative sample of firms.

The paper is organized as follows: Section 2 discusses briefly the apprenticeship system in Switzerland. Section 3 summarizes previous studies and motivates the hypothesis tested in this paper. Section 4 introduces the data and the sample design. In Section 5 (potential) net costs are estimated using a maximum likelihood selection model. Then, the effect of net cost on the number of apprentices is estimated using different types of count data models. Section 6 concludes.

2. The apprenticeship system in Switzerland

The apprenticeship system is the route chosen by about 60 per cent of the Swiss youngsters at upper secondary level. More than 180,000 adolescents are currently embarked on what is called the dual education system, i.e. a training programme combining vocational education at school with training in and work for a company. Almost half of the remaining 40 per cent of young people who complete compulsory education go on to attend grammar school (Gymnasium) to prepare them for university and a more academic career. The remainder either opt for other entirely school-based forms of education or pursue no form of post-compulsory education, placing Switzerland ahead of other Organization for Economic Cooperation and Development (OECD) countries in terms of the percentage of the over-16 population attending any form of non-compulsory schooling.

Vocational training in a dual education programme usually lasts 3–4 years. Dropout rates were fairly low in the past. Based on the responses of training firms in the survey used for this paper, the dropout rates were on average 5 per cent. Apprentices graduate with a diploma recognized throughout Switzerland attesting that the apprentice has a professional qualification. The quality of the training provided in Switzerland, which combines school lessons (1–2 days a week) with on-the-job training in a firm under the supervision of certified staff, is recognized internationally as meeting top standards. International comparisons show, in terms of scholastic and professional qualifications, that Swiss apprentices are more than a match for their upper secondary-level peers attending school full-time (see, for example, Bierhoff and Prais, 1997).

The employment period ends automatically on completion of training. Any extension of the employment period must be negotiated in a separate contract. Switzerland differs in this respect from some other countries where apprentices are protected from dismissal for a period of time after completing their training. Mobility is fairly high among young people who complete their apprenticeship, with only 36 per cent still working at their original training site 1 year on (see Schweri *et al.*, 2003).[1]

3. Previous studies

There have been several studies on the firms' training decision in the last few years, but most of them did not include data on the cost–benefit ratio of apprenticeship training at the firm level. Some of the studies described below used proxy variables for the cost–benefit ratio instead, and others have simply ignored this possibility. On the other hand, the studies that have calculated the cost and benefit of apprenticeship training did not use this information further to analyse its impact on the firms' training decision.

Concerning the training decision of firms, Harhoff and Kane (1997) found for Germany that 'firms are much more willing to train when there are fewer firms around to poach their trainees' (p. 184). They also showed that firms are less likely to train or train less apprentices if the work force in counties, or counties in commuting distance, is large.

Neubäumer and Bellmann (1999) also examined German firms without data on cost and benefit. Their data set contained, however, qualitative assessments by the non-training firms of the factors affecting the training decision. Their findings were that about two-thirds of the firms do not train because they do not meet the legal requirements for offering apprenticeship training. For the remaining third of the non-training companies, an unfavourable cost–benefit ratio was one of the main reasons not to train. Further, the analysis of the training intensity of training firms showed that the firm size is the major factor explaining the ratio of apprentices and total number of employees. Similar to Switzerland, the majority of all apprentices in Germany are trained by small firms with less than 50 employees. Wolter and Schweri (2002) found similar results for the training intensity among training firms in Switzerland, but in addition, the net cost of training also had a significant impact, even if other factors such as the firm size or the training profession were accounted for.[2] This result goes in line with the observation made by Neubäumer and Bellmann (1999) that the training intensity was particularly high in industries where previous studies (see Von Bardeleben *et al*., 1995) had shown low or even negative net costs for apprentices.[3]

Stöger and Winter-Ebmer (2001) found that the number of apprentices trained in Austria declined over time, but they were not able to find explanations for this time trend. As in Germany and Switzerland, they observed that large firms are more likely to train apprentices, but the intensity of training is the highest for small firms.

In a different type of analysis, Fougère and Schwerdt (2002), comparing the French and the German apprenticeship system, analysed the determinants of firms' offer of apprenticeship positions. Using a production function approach, they found evidence that apprentices participate significantly in the production process only in medium-size firms. They conclude from this result that small and large firms train apprentices less for the motive of exploiting the value of their productivity, but rather because of their difficulty to find skilled workers on the external labour market.

Beckmann (2002) finally used a zero-inflated negative binomial model to test the implications of the theory proposed by Acemoglu and Pischke (1998, 1999a, 1999b) for firms from West and East Germany. He defined two regimes, one where firms will never train apprentices, and a second regime, in which firms can both train zero or a positive number of apprentices. He found evidence that a high degree of estimated wage compression encourages firms to train apprentices. His findings apply only for companies for which he assumed positive net costs of apprenticeship training, as he excluded firms from his sample for which he assumed zero or negative net cost of training (companies with less than five employees). Such a reduction of the sample of analysed firms is in our view inappropriate to analyse the overall decision of firms to train apprentices, both in Germany and in Switzerland. Furthermore, with the data used in this study, such assumptions are not necessary as the net cost of training can be included directly in the estimation of the number of apprenticeship positions offered by a firm.

4. Data

The data used here are from two representative surveys conducted in Swiss firms in the year 2001 by the Centre for Research in Economics of Education at the University of Berne and the Swiss

Federal Statistical Office (see Schweri *et al.*, 2003 or Wolter and Schweri, 2002). In one survey, training firms were asked about their cost and benefit of apprenticeship training. In the second survey, non-training firms were asked to fill in the identical questionnaire as training firms, but without the questions directly related to the costs and benefits of training. The original data set contains 2,352 training firms and 2,230 non-training firms, but the public sector has been excluded, because the profit-maximizing principle does not fully apply to such firms. Also excluded are firms that cannot make independent decisions about training because they are part of a larger enterprise. The data set used in this paper embraces 1971 training firms and 1661 non-training firms. Table 1 shows the means and standard deviations of the variables.

The data set is very similar to a German survey that was carried out by the 'Bundesinstitut für Berufsbildung' in Bonn (BIBB) for many years. The questionnaires are specifically designed to calculate the costs and benefits of apprenticeship training for training firms on the one hand and to find out why firms do not train on the other hand; all variables that are seen as important in the literature have been included. This gives us the advantage over other studies, in that we do not need to rely on a reduced set of variables.

For training firms, the net costs of apprenticeship training are different between the cost and benefits. The observed costs were calculated as follows: training costs are the wages of apprentices and the cost for the training personnel, which add up in equal shares to about 90 per cent of total cost. The remaining 10 per cent are cost for material, infrastructure, external courses, and other.

The calculation of training costs shows clearly that the main part of it is driven by wages. Consequently, training cost differences between firms depend on variables that influence either the wage level of apprentices or training personnel. The wages of apprentices are closely related to wage recommendations of professional associations. Hence, the variation of wages within a profession is relatively small, and a substantial part of this can be explained with the size of a firm. Larger firms tend to offer higher wages to apprentices, which is in line with the fact that larger firms offer higher wages to all categories of workers. Analogously to standard wage regressions, the average wages of the training personnel in a firm can be best explained using variables such as firm size, industry, and regional characteristics. The remaining cost for material, infrastructure, and external courses are mainly determined by the training profession.

Table 1. Sample descriptives (N = 3,632; sampling weights have been used)

Variable	**Mean**	**Std. Dev.**
Training firm	0.291	0.454
Number of apprentices	0.675	3.878
Firm size: 1–3	0.331	0.471
Firm size: 4–9	0.403	0.491
Firm size: 10–49	0.224	0.417
Firm size: 50–99	0.023	0.151
Firm size: >100	0.019	0.135
ln (number of skilled workers)	0.313	3.035
Sector: service	0.679	0.467
Industry	0.136	0.343
Construction	0.119	0.324
German part of Switzerland	0.750	0.433
French part of Switzerland	0.222	0.416
Italian part of Switzerland	0.029	0.167
Foreign firm ownership	0.116	0.320
Difficulties in finding qualified labour	0.403	0.491
Reduction in school days	0.113	0.317
Commercial employee	0.177	0.381
Polymechanics technician	0.019	0.136
IT specialist	0.028	0.164
Cook	0.070	0.255
Electromechanics technician	0.020	0.139
Mason	0.025	0.157
Architectural draftsperson	0.030	0.171
Salesperson (2 years)	0.062	0.242
Salesperson (3 years)	0.026	0.159
Auto mechanic	0.020	0.141
Carpenter	0.025	0.155
Office worker	0.033	0.180
Assistant in a doctor's office	0.021	0.145
Structural draftsperson	0.010	0.100
Hairdresser	0.017	0.129
Automation technician	0.004	0.063
Electronics technician	0.004	0.064
Other professions	0.409	0.492

The benefits are calculated by the type of work the apprentices perform. It is broken down into production activities that would otherwise be performed by unskilled workers or skilled workers. Although we can assume in the first case that the apprentice's performance has the same value as that of an unskilled worker, the value of the apprentice's performance for the second case is compared with that of a fully trained skilled worker. Much of the variation in the benefits of apprenticeship training can be explained with the determinants of both the apprentices wages and the wages of skilled workers, and the profession in which an apprentice is trained.[4]

Given the importance of training professions for the net cost calculation, it is a major advantage of the data set used in this study that the survey was carried out profession specific. It should be noted that the training profession has only been included in the study of Beicht *et al.* (2004) and the data set used in this study; all earlier studies (see below) of the cost and benefits of apprenticeship training had no information on the trained profession.

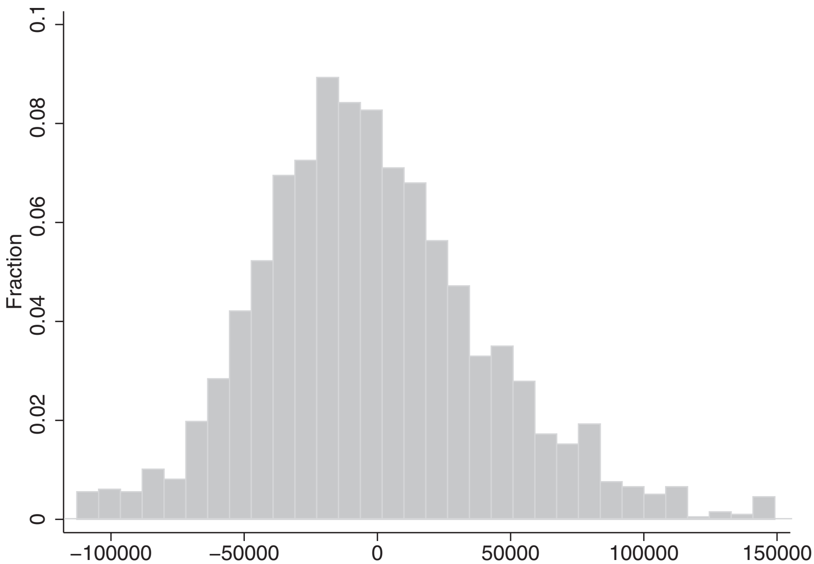
Other variables potentially influencing cost and benefit of apprenticeship training used in the literature, such as process innovation, product innovation, current revenue, modernity of technical equipment, as well as expectations of future revenue (see Franz *et al.*, 2000 or Beckmann, 2002), were also tested in our analysis. However, we did not find any significant impact either on the training decision or on the net costs of training.[5]

Figure 1 shows the distribution of net cost for the 1971 training firms. The average is negative (benefits exceed cost) at CHF –6,174. Net costs are negative for 60 per cent of all training firms and depend largely on the training profession, firm size, industry, and regional characteristics for the reasons given above. The findings confirm earlier studies that directly measured the costs of apprenticeship training, e.g. Von Bardeleben *et al.* (1995) for Germany, Lassnigg and Steiner (1997) for Austria, and Hanhart and Schulz (1998) for Switzerland.[6]

5. Econometric models and empirical analysis

In order to estimate the effect of net cost of training on the training decision, we require the potential net cost for currently non-training firms, which we naturally cannot observe. Hence, we need to estimate these counterfactual costs, and the first subsection shows how to do that. In the next subsection, we present different count data models in order to estimate the effect of predicted net cost on the number of apprentices employed in a firm. A particular feature of the data is the high proportion of zeros, i.e. the large fraction of non-training firms. These observations represent a corner solution. Therefore, we concentrate on count data models

Figure 1. Distribution of net cost for training firms



for corner solutions, so-called 'hurdle models'. A useful property of these models in the present context is that they allow us to distinguish between cost elasticities at the extensive and intensive margins. The third subsection presents the results.

5.1 Estimation of potential net cost

The net costs of training are observed only for training firms. To use these costs to estimate the potential costs of training for currently non-training firms, we have to realize that training firms are not a random sample from the universe of all firms, and that training and non-training firms may differ systematically in observable and unobservable characteristics. In order to account for observable differences, we use a linear regression framework according to which

$$y_{1j} = x'_{1j}\beta_1 + \varepsilon_{1j} \quad j = 1, \ldots, 3{,}632, \qquad [1]$$

where y_{1j} is net cost and x_{1j} is a vector containing variables such as firm size, number of skilled workers, industry, apprenticeship profession, region, ownership of the firm, and a variable indicating whether a firm would like to reduce the time an apprentice spends in vocational school.[7]

We do not estimate a log-linear model because net cost can be positive as well as negative. If $\hat{\beta}_1$ is a consistent estimator for β_1, we can predict the net cost for non-training firms consistently as $\hat{y}_{1j}^{nt} = x_{1j}^{nt\prime}\hat{\beta}_1$.

However, direct estimation of [1] by ordinary least squares using training firms only leads to selection bias as long as training and non-training firms differ in unobservables as well observables. Let

$$y_{2j} = \begin{cases} 1 & \text{if firm } j \text{ trains} \\ 0 & \text{if firm } j \text{ does not train} \end{cases}. \qquad [2]$$

Selection bias arises if $E(\varepsilon_{1j}|y_{2j} = 1) \neq E(\varepsilon_{1j}|y_{2j} = 0)$. We follow here a standard Heckman approach with a latent continuous variable for the training decision:

$$y_{2j} = 1[x_{2j}'\beta_2 + \varepsilon_{2j} < 0]. \qquad [3]$$

This is a reduced form version of the firms' profit maximization condition, according to which firms train as long as the marginal benefit of doing so is positive. The benefit of training has two components. First, an immediate benefit arises if the net costs of training are negative. Second, a benefit can arise in later periods, for example, because the firm needs skilled workers and it may be difficult to recruit such workers in the outside labour market. This reasoning implies that x_{2j} should include all determinants of net cost x_{1j}. These variables obviously affect the decision to train, although not necessarily through their influence on net cost only. They may also have an indirect effect, by affecting other aspects of the training decision. In addition, x_{2j} includes a variable measuring the tightness of the labour market for skilled workers (the exclusion restriction). This variable, a dummy indicating whether a firm has difficulties finding skilled workers on the labour market, does not influence the net cost of the apprenticeship training but it is expected to have a significant impact on the firm's decision to train when the net cost of the apprenticeship training alone would be positive. The variable on labour market tightness has also been used to explain the training behaviour of firms in a number of other studies, see, for example, Askilden and Nilsen

(2005) for apprenticeship programmes in Norway or Majumdar (2007) for a recent example using US data.

If we moreover assume that (ε_{1j}, ε_{2j}) are bivariate normal distributed with zero mean, covariance σ_{12}, and variances σ_1^2 and one, respectively, we can estimate all model parameters by maximum likelihood.

The contribution of firm j to the log-likelihood function is

$$L_j = (1 - y_{2j})\ln[1 - \Phi(x'_{2j}\beta_2)] + y_{2j}\left\{\ln\Phi\left[\frac{x'_{2j}\beta_2 + \sigma_{12}\sigma_1^{-2}(y_{1j} - x'_{1j}\beta_1)}{\sqrt{1 - \sigma_{12}^2\sigma_1^{-2}}}\right]\right.$$
$$\left. + \ln\phi\left(\frac{y_{1j} - x'_{1j}\beta_1}{\sigma_1}\right) - \ln(\sigma_1)\right\}$$

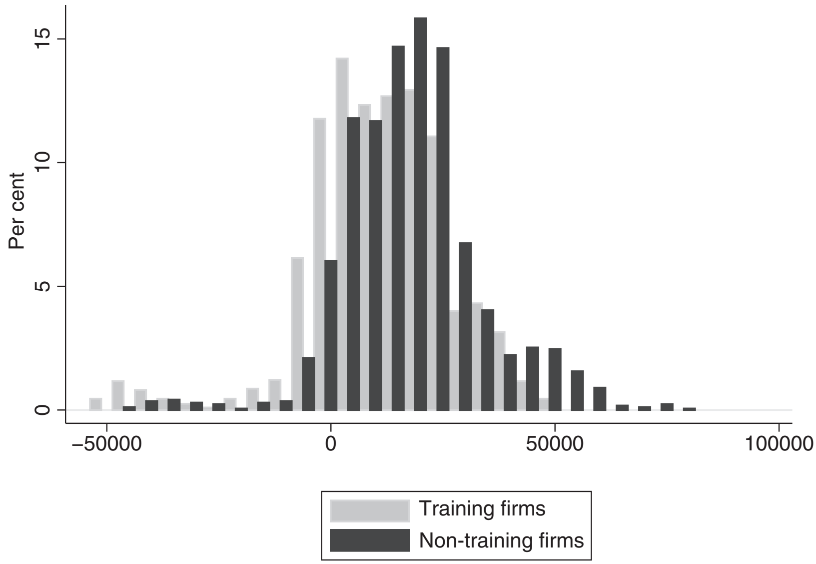
(see Wooldridge, 2002).[8] The estimation results are given in Table 2. The estimated $\rho = \sigma_{12}/\sigma_1$ is positive. The point estimate is 0.5, and a

Table 2. Maximum likelihood estimation selection model, dependent variable: net cost of training

	Net cost		Training yes/no	
	Coeff.	**Std. err.**	**Coeff.**	**Std. err.**
Industrial sector	–1,510.6	(2,687.3)	0.145[a]	(0.066)
Construction sector	–7,799.4[a]	(3,579.8)	0.329[a]	(0.087)
French part of Switzerland	1,733.4	(2,580.1)	–0.240[a]	(0.060)
Italian part of Switzerland	8,479.1	(5,750.7)	0.044	(0.144)
Foreign-owned company	5,617.6	(3,212.0)	–0.615[a]	(0.074)
Firm size: 4–9 employees	–3,631.6	(3,561.6)	0.168[a]	(0.069)
Firm size: 10–49 employees	–6,742.8	(4,148.2)	0.362[a]	(0.085)
Firm size: 50–100 employees	–8,341.6	(5,029.0)	0.862[a]	(0.109)
Firm size: >100 employees	–4,287.8	(5,499.7)	1.104[a]	(0.121)
Reduction in school days	8,327.6[a]	(3,753.6)	–0.368[a]	(0.082)
Number of skilled workers (ln)	–2,498.7[a]	(1,095.1)	0.317[a]	(0.031)
Difficulties in finding qualified labour			0.303[a]	(0.049)
ρ	0.503	(0.079)		
σ_1	42,580.8	(1,152.2)		
Log-likelihood	–25,551.2			
Observations	1,971		3,632	

Notes: [a] Effect is significant at the 5 per cent level. The reference site is a Swiss-owned company located in the German-speaking part of Switzerland, which has more than 100 employees. The model includes in addition a constant and 17 indicator variables for the training profession.

Figure 2. Overlapping histograms for predicted net cost of training firms (light colour) and non-training firms (dark colour)



Wald test for the null hypothesis $H_0 : \rho = 0$ has a P-value of 0.0013, so that the null is rejected. This implies that the higher the unobserved net cost component ε_{1j}, the lower is the probability of training. Training firms self-select based on absolute cost advantage.

If we use the selectivity-corrected net cost estimates in columns 2 and 3 of Table 2, we find that the predicted net cost of an average apprenticeship programme of a randomly drawn firm in Switzerland amounts to about CHF 13,608[9] (Table 2). Of course, there is considerable variation across firms, as seen in Figure 2, which shows a histogram of the predicted net cost of training and non-training firms. The predictions are now based on [1] and thus conditional on x_{1j} but unconditional on the training decision. As expected, predicted net costs tend to be higher for non-training firms than for training firms — the averages are CHF 18,540 and CHF 9,453, respectively. However, there is considerable overlap in the two distributions. Of all non-training firms (about 8,090 firms), 5.45 per cent are predicted to have negative net cost and still do not train. At the point of time of the survey some 5,000 offered training posts were not filled because of a lack of either any, or any appropriate, applicants. Therefore, we conjecture that more than half of these non-training firms with negative net cost did not train for this reason. According to our model, the rest of these firms must have a positive unobserved cost component.

In the following analysis, the predicted net costs — unconditional on the training decision — are used as an explanatory variable in a structural model in order to estimate the cost elasticity.

Because the dependent variable, the number of apprentices, is a non-negative integer, we discuss first appropriate econometric models for such counts.

5.2 Count data models for the number of apprentices employed by a firm

Let $n_j = 0, 1, 2, \ldots$ denote the number of apprentices employed by firm j and $\hat{y}_{1j}$ the predicted net cost that can be positive or negative. The main objects of interest are the cost elasticity at the extensive margin

$$\eta_1 = \frac{\partial P(n_j > 0)}{\partial \hat{y}_{1j}} \frac{\hat{y}_{1j}}{P(n_j > 0)},$$

and the cost elasticity at the intensive margin

$$\eta_2 = \frac{\partial E(n_j | n_j > 0)}{\partial \hat{y}_{1j}} \frac{\hat{y}_{1j}}{E(n_j | n_j > 0)}.$$

As $E(n_j) = P(n_j > 0)E(n_j|n_j > 0)$ it follows that the overall cost elasticity equals $\eta_1 + \eta_2$. For example, in the Poisson regression model, where $P(n_j|\lambda_j) = \exp(-\lambda_j) n_j^{\lambda_j} / n_j!$ and $\lambda_j = \exp(x_j'\delta + \gamma \hat{y}_{1j})$, we obtain that

$$\eta_1 = \left[\frac{\lambda_j \exp(-\lambda_j)}{1 - \exp(-\lambda_j)} \right] \gamma \hat{y}_{1j} \quad \text{and} \quad \eta_2 = \left[1 - \frac{\lambda_j \exp(-\lambda_j)}{1 - \exp(-\lambda_j)} \right] \gamma \hat{y}_{1j}.$$

Thus, $\eta_1 = \eta_2 = 0$ if and only if $\gamma = 0$. Obviously, this model does not allow for testing separately for the statistical significance of intrinsic and extrinsic cost elasticities — as we would like to. An econometric tool that allows one to perform such separate tests is provided by the class of so-called count data hurdle models (Mullahy, 1986; Winkelmann, 2004).

A count data hurdle model combines a binary model for the training decision with a truncated-at-zero count data model for the number of apprentices employed by training firms. The standard formulation is the Poisson-hurdle (PH) model where the binary model is of a complementary log–log form. The advantage of the PH model is that it nests the simple Poisson model, which thus can be easily tested against the more general hurdle model. We present PH model as a benchmark, and in order to assess the robustness of the main conclusion with respect to the parametric assumptions we make.

For a number of reasons, we suspect that an alternative, recently developed hurdle model, the probit Poisson log-normal (PPLN) model (Winkelmann, 2004), is more appropriate in the present context. First, it retains the probit structure of the binary training equation [3]. The main difference is that we now include the predicted net cost of training $\hat{y}_{1j}$ as an additional regressor. The variable 'reduction in vocational school time is desired (yes/no)' is included in the cost equation but excluded in the demand equation.

Second, the PPLN model allows for unobserved heterogeneity in the positive part of the hurdle model. In previous research, it has been found that the alternative approach of using the negative binomial model, based on mixing over a gamma distribution, is usually inferior to mixing over a log-normal distribution (Winkelmann, 2003). In particular, assume that

$$\lambda_j = \exp\left(x_{3j}'\delta_2 + \gamma_2 \hat{y}_{1j} + u_j\right), \tag{4}$$

where $u_j \sim \text{Normal}(0,\ \sigma_u^2)$ independently of x_{3j} and $\hat{y}_{1j}$. If $n_j|u_j$ is Poisson distributed, then the unconditional model n_j has the Poisson log-normal form. Although closed form expressions are not available, the probabilities can be evaluated by Gauss–Hermite quadrature. If the Poisson log-normal model is combined with a probit hurdle as in [3], we obtain the PPLN model with probability function

$$f(n_j) = [1 - \Phi(x_{3j}'\delta_1 + \gamma_1 \hat{y}_{1j})]^{1(n_j=0)} \Bigg\{ \Phi(x_{3j}'\delta_1 + \gamma_1 \hat{y}_{1j}) \times \int_{-\infty}^{\infty} \frac{\exp[-\lambda_j(u_j)][\lambda_j(u_j)]^{n_j}}{[1-\exp(-\lambda_j(u_j))]n_j!} \frac{1}{\sigma_u} \phi(u_j/\sigma_u)\mathrm{d}u_j \Bigg\}^{1(n_j>0)}, \tag{5}$$

where ϕ denoted the standard normal density. The parameters δ_1, δ_2, γ_1, γ_2, and σ_u^2 can be jointly estimated by maximum likelihood. A useful property of all standard hurdle models is that the log-likelihood function has two additive parts $L_1(\delta_1, \gamma_1)$ and $L_2(\delta_2, \gamma_2, \sigma_u^2)$ that can be maximized separately with respect to the corresponding parameters. In the PPLN model described here, $\hat{\delta}_1$ and $\hat{\gamma}_1$ are obtained from estimating a probit model with all data whereas the remaining parameters are obtained by maximizing a truncated-at-zero Poisson log-normal model using observations on training firms only.

We can now return to the question of the elasticities at the extensive and intensive margins. For the extensive margin, the cost elasticity is readily calculated as

$$\eta_1 = \frac{\phi(x'_{3j}\delta_1 + \gamma_1\hat{y}_{1j})}{\Phi(x'_{3j}\delta_1 + \gamma_1\hat{y}_{1j})}\gamma_1\hat{y}_{1j}.$$

The cost elasticity for training firms cannot be calculated in closed form but it needs to be evaluated numerically.

5.3 Results

Tables 3 and 4 provide the estimation results for the PH and the PPLN model. We report robust standard errors in all instances. Several tests are possible. First, a comparison of the PH and the simple Poisson model shows that the null hypothesis of equal coefficients in the hurdle and in the positive part is clearly rejected. Under the Poisson restriction, the likelihood ratio test statistic of 754.44 is a realization from a $\chi^2(30)$ distribution with critical value 43.77. The PH and the PPLN model are not nested. However, the substantial drop in log-likelihood from −8,522.7 to −5,262.2 with just one additional parameter indicates a clear preference for the PPLN model, based, for example, on the Akaike Information Criterion. Clearly, unobserved heterogeneity is an important feature of these data. This view receives support from formal tests between non-nested models, such as the Vuong test (Vuong, 1989). We also estimated a couple of other hurdle models available in Stata, such as a logit-negative binomial model and a complementary log–log-negative binomial model. We found that the PPLN model proposed here dominates both of them. Moreover, we found in all instances the same substantive result with regard to the cost elasticities at the intensive and extensive margin, namely a significant negative effect

Table 3. Demand for apprentices — PH model

	Training yes/no		No. apprentices 1+	
	Coeff.	**Std. err.**	**Coeff.**	**Std. err.**
Industrial sector	0.147	(0.073)	–0.158	(0.162)
Construction sector	0.090	(0.124)	–0.263	(0.213)
French part of Switzerland	–0.232	(0.073)	–0.808	(0.194)
Italian part of Switzerland	0.346	(0.194)	–0.264	(0.325)
Foreign-owned firm	–0.389	(0.099)	–0.003	(0.178)
Firm size: 4–9 employees	0.115	(0.096)	–0.224	(0.174)
Firm size: 10–49 employees	0.175	(0.125)	–0.330	(0.297)
Firm size: 50–100 employees	0.587	(0.148)	–0.413	(0.372)
Firm size: >100 employees	0.928	(0.137)	–0.035	(0.290)
Number of skilled workers (ln)	0.206	(0.040)	0.685	(0.064)
Difficulties in finding qualified workers	0.330	(0.055)	0.128	(0.122)
Net cost of training (in thousands)	–0.046	(0.011)	–0.025	(0.025)
Log-likelihood	–1,876.7		–6,646.6	
Observations	3,632		1,971	

Notes: Robust standard errors in parentheses. The model includes in addition a constant and 17 indicator variables for the training profession.

at the extensive margin, and no significant effect at the intensive margin (see Table 3 for the PH results).

Table 4 gives the details for the regression coefficients in the PPLN model. The first column gives the probit coefficients, whereas the third column shows the coefficients for the truncated-at-zero Poisson log-normal model. The main parameter of interest is the structural effect of net cost. The regression parameter is negative and statistically significant in the hurdle part but practically zero for the positives. This means that an increase in net costs reduces the probability of offering (any) training, but such an increase is unrelated to the number of apprentices a firm trains, once it has decided to train. A potential economic explanation for this pattern is that small firms who train a small number of apprentices have in general lower net costs of training than large firms training many apprentices. Although training is lucrative for small firms, the number of apprentices they employ is limited by an upper bound that depends on the number of skilled workers able to train apprentices. If they would extend their offer of apprenticeship places beyond that limit, the additional need for trainers (and additional infrastructure) would result in positive marginal net costs for an additional

Table 4. Demand for apprentices — PPLN hurdle model

	Training yes/no		**No. apprentices 1+**	
	Coeff.	**Std. err.**	**Coeff.**	**Std. err.**
Industrial sector	0.072	(0.069)	–0.061	(0.063)
Construction sector	–0.013	(0.114)	–0.111	(0.125)
French part of Switzerland	–0.150	(0.065)	–0.465	(0.068)
Italian part of Switzerland	0.374	(0.174)	–0.372	(0.162)
Foreign-owned firm	–0.373	(0.090)	–0.044	(0.085)
Firm size: 4–9 employees	0.013	(0.077)	0.081	(0.135)
Firm size: 10–49 employees	0.063	(0.107)	0.346	(0.154)
Firm size: 50–100 employees	0.500	(0.135)	0.496	(0.168)
Firm size: >100 employees	0.919	(0.130)	0.823	(0.154)
Number of skilled workers (ln)	0.208	(0.040)	0.522	(0.035)
Difficulties in finding qualified workers	0.291	(0.050)	0.080	(0.049)
Net cost of training (in thousands)	–0.044	(0.009)	0.012	(0.108)
ln(σ)			–0.589	(0.036)
Log-likelihood	–1,871.6		–3,390.6	
Observations	3,632		1,971	

Notes: Robust standard errors in parentheses. Model includes in addition a constant and 17 indicator variables for the training profession.

apprentice. Large firms with positive net costs of training are not sensitive to marginal variations, because the number of apprentices is determined by the number of future vacancies for skilled workers. In this situation, a reduction in the net costs of training would only translate into an additional number of apprenticeship posts if the marginal net costs of training were smaller than the marginal benefit for the period after the apprenticeship. The latter depends on the probability that the firm is able to offer its former apprentice a job. This together leads to the result that once a firm has decided to train apprentices, the number of apprentices does not depend on marginal variations in the net costs of training. The effect of net costs is therefore entirely restricted to the extensive margin.

To put a quantitative meaning on the estimated coefficient of –0.044, we can compute the elasticity at the average values of the regressors. In this case, $\hat{\eta}_1 = -0.45$ — a 1 per cent increase in the net cost reduces the probability of training by 0.45 per cent. Obviously, these marginal elasticities should not be extrapolated too far, as the probit model is highly non-linear. Alternatively, we can consider absolute percentage point changes in the probability of training.

For example, starting again from average values, an increase in net training cost by 1,000 will reduce the probability of training by 1.8 percentage points. Similarly, an increase in net training cost by one standard deviation reduces the probability of training by 27.4 percentage points.

In order to summarize the main result of the analysis, we find that the elasticity at the intensive margin is zero, whereas the estimated elasticity at the extensive margin amounts to –0.45, an economically substantial effect. We conclude by mentioning some of the other results. Firm size, measured by the number of employees, and the number of skilled workers have the expected positive effects both on the probability of training and on the number of apprentices among training firms. Foreign-owned firms are less likely to train, as are firms in the French-speaking part of Switzerland. The training firms in the French part in addition train fewer apprentices — for given firm size, net cost, etc. — than otherwise similar firms in the German-speaking part of Switzerland. These results largely confirm those from previous studies.

5.4 A simulation for subsidies

In the year 2000, the year of reference of our survey, about 74,500 new apprentices were employed by Swiss firms, but not every school leaver interested in apprenticeship training was able to find a training position. In subsequent years, the gap between the number of school leavers interested in apprenticeship training and the number of available apprenticeship positions increased steadily due to the unfavourable general economic climate. Political initiatives tried to fight the imbalance in the apprenticeship market with the idea of subsidizing firms for their training. Taking into account our findings, the subsidy should be given to non-training firms, where it would have a large effect, but not to training firms, where it would have no effect at all. A cost subsidy to a non-training firm would reduce net cost of training and thereby increase the probability to offer a training position. The question remains how much such a programme would cost.

In the year 2004, according to the 'Lehrstellenbarometer (2004)', some 8,000 interested school leavers did not succeed in finding a training position. With an average of 1.04 newly created apprenticeship posts per training firm, one would need to attract 7,666 new training firms, an increase in the proportion of training firms by 3.66 percentage points. As shown in the previous section, the training ratio increases by 1.8 percentage points for each reduction in cost by CHF 1,000. Thus, the targeted increase by 3.66 percentage points requires a subsidy of CHF 1,949 per apprentice.

Of course, it would be difficult in practice to assess whether a given firm would have offered no training post in the absence of a subsidy, and thus to discriminate between firms who do already train apprentices and new firms when deciding whom to offer the training subsidies. In the spirit of our model, one would need to have detailed information on the net cost of training for that firm. However, firms can be expected to respond strategically and overstate their true cost once such a scheme would be in place. Therefore, all known political initiatives demanding subsidies for training start with the idea that each training post would be subsidized, and not only the 'additional' ones. This windfall gain for 'old' training companies creates as a consequence much higher total costs of the subsidy than a regime that could be targeted at the new companies only. As a consequence, the subsidies alone would amount to more than CHF 17,000 per newly created apprenticeship position.

The amount depends on the number of newly created apprenticeship places but is basically driven by the total number of training firms. To illustrate this, a scenario with an additional 1,000 and another scenario with 10,000 additional new apprenticeship places are shown in Table 5. To the costs of subsidies, the costs of school for an apprentice must be added, which amount on average to CHF 15,000 per year. The total cost of education (subsidy included) of an additionally trained apprentice would be more than CHF 60,000. For purposes of comparison, the total costs for the highest form of a full-time general education in Switzerland (Gymnasium) are on average CHF 58,500 for a 3-year period. Taking into account that the administration of such subsidies would generate additional costs of 10–20 per cent of the total amount of subsidies, the 'artificial' creation of new apprenticeship places through subsidies seems questionable from a cost point of view.

Table 5. Subsidy simulation

New apprenticeship positions	Total cost of Subsidies (in million CHF)	Subsidy per newly created apprenticeship position (in CHF)	Total cost per apprenticeship position (in CHF)
8,000	139.7	17,463	62,463
1,000	15.8	15,758	60,758
10,000	179.5	17,950	62,950

6. Concluding remarks

The paper has made both a methodological and a substantive contribution. On the methodological side, it is the first time that the training decision is modelled in a structural framework. In this empirical framework, the main parameter of interest, the cost elasticity, can be identified from observing cost data for training firms alone. The problem that costs cannot be observed for non-training firms is overcome by using predictions from a selectivity-corrected cost equation. A hurdle count data model is then used to estimate the structural equation, so that cost elasticities at the extensive and intensive margins can be estimated separately. To estimate the model, we used a unique firm-level data set that includes training and currently non-training firms and provides detailed cost information for training firms.

The results are quite striking. We find that the cost elasticity at the extensive margin, i.e. with regards to the probability whether to train or not to train, amounts to –0.45, an economically substantial effect. In contrast, the cost elasticity at the intensive margin, i.e. for the number of apprentices among training firms, is zero.

We close with the substantive conclusion that in order to increase the number of apprenticeship positions offered each year, one would need to direct subsidies at non-training firms and exclude training firms. In this case, the required subsidy would be quite modest. If, however, such a discrimination is politically and practically infeasible, the costs for creating additional apprenticeship positions would be prohibitively high.

Notes

[1] In Germany, the corresponding figure is closer to 70 per cent (see Winkelmann, 1996; Euwals and Winkelmann, 2002).

[2] Niederalt *et al.* (2001) found similar results for a small sample of 27 Bavarian companies for which they had calculated the cost of apprenticeship training.

[3] An assessment of the German situation was repeated in an analysis made by Dietrich (2000).

[4] For more details on the cost and benefit model used in this study see Wolter *et al.* (2006) or Schweri *et al.* (2003).

[5] As we do not have a panel data set, we cannot observe the effects of real economic and demographic variables over time, but we can proxy them by the firm's expectation about the future developments.

[6] So far, detailed and representative studies that directly measure the costs and benefits of apprenticeship training have — to our knowledge — only been carried out in German-speaking countries. However, the OECD has started a collaborative project in its statistical network to establish guidelines on how to measure the cost of vocational training.

[7] The last variable will be used later on to identify the cost elasticity in a 'structural' model. It is assumed to affect the number of apprentices only through net cost.

[8] The null hypothesis of normality could not be rejected using a diagnostic test proposed by Pagan and Vella (1989, p. S51).

[9] These results differ from the results in Wolter *et al.* (2006) because weights were used in the earlier paper. The use of weights is not required to obtain consistent estimates of the regression parameters although differences in small samples occur. From a qualitative point of view, the results remain the same with or without the use of weights. The weights have been dropped as some of the following count data models have not been implemented for use with weights.

References

Acemoglu D. and Pischke J.-S. (1998) 'Why Do Firms Train? Theory and Evidence', *Quarterly Journal of Economics* 113: 79–119.

Acemoglu D. and Pischke J.-S. (1999a) 'Beyond Becker. Training in Imperfect Labour Markets', *Economic Journal* 109: F112–F142.

Acemoglu D. and Pischke J.-S. (1999b) 'The Structure of Wages and Investment in General Training', *Journal of Political Economy* 107: 539–572.

Askilden J. E. and Nilsen O. A. (2005) 'Apprentices and Young Workers: A Study of the Norwegian Youth Labour Market', *Scottish Journal of Political Economy* 52(1): 1–17.

Beckmann M. (2002) 'Wage Compression and Firm-sponsored Training in Germany: Empirical Evidence for the Acemoglu–Pischke Model from a Zero-inflated Count Data Model', *Applied Economics Quarterly (Konjunkturpolitik)* 48(3–4): 368–387.

Beicht U. *et al.* (2004) *Kosten und Nutzen der betrieblichen Berufsausbildung in Deutschland*, Berichte zur beruflichen Bildung, Heft 264, Bonn: Bundesinstitut für Berufsbildung.

Bierhoff H. and Prais S. J. (1997) *From School to Productive Work. Britain and Switzerland Compared*. The National Institute of Economic and Social Research. Cambridge: Cambridge University Press.

Dietrich H. (2000) 'Betriebliches Ausbildungsverhalten im Kontext der betrieblichen Altersstruktur' in George R. and Struck O. (eds.) *Generationenaustausch im Unternehmen*, München: Rainer Hampp Verlag: 159–176.

Euwals R. and Winkelmann R. (2002) 'Mobility after Apprenticeship — Evidence from Register Data', *Applied Economics Quarterly (Konjunkturpolitik)* 48(3–4): 256–278.

Fougère D. and Schwerdt W. (2002) 'Are Apprentices Productive?', *Applied Economics Quarterly (Konjunkturpolitik)* 48(3–4): 317–346.

Franz W. *et al.* (2000) *Die betriebliche Ausbildungsbereitschaft im technologischen und demographischen Wandel*, Schriftenreihe des ZEW, Band 46, Baden-Baden: Nomos Verlagsgesellschaft.

Hanhart S. and Schulz H.-R. (1998) *Lehrlingsausbildung in der Schwweiz*, Chur: Rüegger Verlag.

Harhoff D. and Kane T. J. (1997) 'Is the German Apprenticeship System a Panacea for the U.S. Labor Market?', *Journal of Population Economics* 100: 1147–1158.

Lassnigg L. and Steiner P. (1997) 'Die betrieblichen Kosten der Lehrlingsausbildung, Kammer für Arbeiter und Angestellte für Wien', Materialien zu Wirtschaft und Gesellschaft Nr. 67.

Lehrstellenbarometer (2004) 'Lehrstellenbarometer August 2004, Link Institut Luzern im Auftrag des Bundesamtes für Berufsbildung und Technologie BBT', http://www.bbt.admin.ch/berufsbi/projekte/barometer/archiv/d/kurzbericht_aug_04_d.pdf.

Majumdar S. (2007) 'Market Conditions and Worker Training: How Does It Affect and Whom?', *Labour Economics* 14: 1–23.

Mullahy J. (1986) 'Specification and Testing in Some Modified Count Data Models', *Journal of Econometrics* 33: 341–365.

Neubäumer R. and Bellmann L. (1999) 'Ausbildungsintensität und Ausbildungsbeteiligung von Betrieben: Theoretische Erklärungen und empirische Ergebnisse auf der Basis des IABBetriebspanels 1997' in Beer D., Frick B., Neubäumer R. and Sesselmeier W. (eds.) *Die wirtschaftlichen Folgen von Aus- und Weiterbildung*, München und Mering: Rainer Hampp Verlag: 9–41.

Niederalt M. *et al.* (2001) 'Betriebliches Ausbildungsverhalten zwischen Kosten-Nutzen-Kalkül und gesellschaftlicher Verantwortung', Friedrich-Alexander-Universität Erlangen-Nürnberg, Diskussionspapiere No. 7.

Pagan A. and Vella F. (1989) 'Diagnostic Tests for Models Based on Individual Data: A Survey', *Journal of Applied Econometrics* 4: S29–S59.

Schweri, J. *et al.* (2003) *Kosten und Nutzen der Lehrlingsausbildung aus der Sicht Schweizer Betriebe*, Beiträge zur Bildungsökonomie, Band 2, Chur und Zürich: Rüegger Verlag.

Stöger K. and Winter-Ebmer R. (2001) 'Lehrlingsausbildung in Osterreich: Welche Betriebe bilden Lehrlinge aus?', Johannes Kepler Universität Linz, Arbeitspapier Nr. 0110.

Von Bardeleben R. *et al.* (1995) 'Betriebliche Kosten und Nutzen der Ausbildung: Reprsentative Ergebnisse aus Industrie, Handel und Handwerk', Berichte zur beruflichen Bildung, No. 187, Bielefeld.

Vuong Q. H. (1989) 'Likelihood Ratio Tests for Model Selection and Non-nested Hypotheses', *Econometrica* 57: 307–333.

Winkelmann R. (1996) 'Employment Prospects and Skill Acquisition of Apprenticeship-trained Workers in Germany', *Industrial and Labor Relations Review* 49(4): 658–672.

Winkelmann R. (2003) *Econometrics Analysis of Count Data*, 3rd edn, Heidelberg: Springer.

Winkelmann R. (2004) 'Health Care Reform and the Number of Doctor Visits — An Econometric Analysis', *Journal of Applied Econometrics* 19: 455–472.

Wolter S. C. and Schweri J. (2002) 'The Cost and Benefit of Apprenticeship Training: the Swiss Case', *Applied Economics Quarterly (Konjunkturpolitik)* 48(3–4): 347–367.

Wolter S. C., Mühlemann S. and Schweri J. (2006) 'Why Some Firms Train Apprentices and Many Others Do Not', *German Economic Review* 7(3): 249–264.

Wooldridge J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, Cambridge, MA: The MIT Press.