Hiring Costs and the Firm's Supply of Training*

Marc Blatter
University of Bern
Samuel Muehlemann
University of Bern & IZA Bonn
Samuel Schenker
University of Bern

January 23, 2009

Abstract

This paper analyzes the effects of hiring costs for skilled workers on the firm's supply of training. Increasing marginal hiring costs make it increasingly expensive for firms to hire skilled workers on the external labor market. However, internal training can be an alternative to recruiting skilled workers externally. Firms may find it profitable to invest in costly training in order to reduce hiring costs. Our empirical results provide evidence that the firm's supply of training in fact depends on the firm's level of hiring costs.

JEL Classification: J23, J24, J32, J42
Keywords: Hiring costs, monopsony, supply of training

*The study is based on two surveys from the years 2000 and 2004. The first is financed by the Commission for Technology and Innovation (CTI credit 4289.1 BFS and 5630.1 BFS), the second by the Swiss Federal Office for Technology and Innovation. We are grateful to Stefan C. Wolter for helpful comments and suggestions. The usual disclaimer holds.

1 Introduction

To recruit skilled workers, firms incur substantial hiring costs. First, they have to post vacancies and process interviews with job-applicants. Second, it takes time for a newly hired worker to reach full productivity in the new job. In short, the firm's hiring costs include both recruitment and adaption costs. However, firms have the possibility to avoid such costs. Instead of hiring all workers on the external labor market, firms can train workers internally until they reach the desired skill level. There are many types of firm training, e.g., on- and off-the-job training, trainee programs, internships or apprenticeship programs.

The effect of hiring costs on the supply of training by firms can be applied to all the types of training mentioned above. However, to make our analysis more tractable, we focus on apprenticeship training, which is particularly widespread in Europe. In Germany and Switzerland, more than half of a cohort of school-leavers choose this type of education. Firms hire young people as apprentices after compulsory schooling. After completion of training, firms may retain them as skilled workers. Obviously, the decision to train apprentices also depends on the costs of training. Training apprentices might be a profitable strategy for a firm to satisfy its labor demand if the costs of training an apprentice are below the costs of hiring a skilled worker on the external labor market. In this sense, the benefits of training apprentices include reduced marginal hiring costs for skilled workers, given that a firm is able to retain a former trainee as a skilled worker after training.

A standard assumption in a model of perfect competition in the labor market is that firms take wages for skilled workers as given. However, there is a large recent literature that provides evidence in favor of monopsonistic labor markets. We carry out our analysis in the framework of a generalized model of monopsony (see Manning (2006)), in which firms set wages. This has several implications. First, by offering a higher wage, firms can reduce labor turnover. Hence, a firm that offers higher wages has to recruit fewer workers to keep employment constant at a certain level. This also implies lower marginal hiring costs, since they depend positively on the number of recruited workers in a given time period. Second, if a firm offers a higher wage, it is more attractive to workers looking for a job. This in turn reduces hiring costs because the firm needs to spend less money to find an appropriate match. We extend the monopsony model such that firms can adjust their labor force not only by recruiting skilled workers on the external labor market, but also by training apprentices within the firm and retaining them as skilled workers after completion of training.

Our results show that an increase in hiring costs leads to a significantly higher supply of training positions by a firm. This is due to the fact that a firm can satisfy its demand for skilled workers by employing former trainees as skilled workers after training. Hence, a firm needs to recruit fewer skilled workers from the external labor market and can thereby avoid the corresponding hiring costs.

The paper is organized as follows. First, we give a brief overview of the related literature. Section 3 outlines the theoretical model of the firm's supply of training. Section 4 provides some important facts about the vocational training system in Switzerland, which we use to illustrate the effects of hiring costs on the firm's supply of training. Section 5 describes the data. Section 6 contains the econometric modeling and the empirical results. Section 7 concludes.

2 Related Literature

Stevens (1994) provides an investment model for the supply of training, in which the employer's return is given by reduced recruitment costs for skilled workers. However, due to the lack of data on both hiring costs and net costs of training, she has to make strong assumptions. For instance, wage data serve as a proxy for the training costs and a variable indicating a shortage of skilled labor is used as a proxy for hiring costs.

Our main interest is the effect of hiring costs on the firm's supply of training. The only studies known to the authors that use direct data on the firms' hiring costs are Manning (2006), Kramarz and Michaud (2004), and Abowd and Kramarz (2003). However, in the British data used by Manning

(2006), hiring costs are only captured as an interval variable. The studies by Kramarz and Michaud (2004) and Abowd and Kramarz (2003) have the limitation that hiring costs do not contain the costs associated with the adaptation of newly hired workers, which may be a substantial part of hiring costs.

Other studies analyze the relation between the costs and the supply of training, but without considering hiring costs. Muehlemann et al. (2007) estimate the firm's supply of training using firm-level data with detailed costs of training. They find that expected costs are an important determinant of the firm's decision to train apprentices. Wolter et al. (2006) show that expected training costs are significantly higher for non-training firms than for firms that do offer training.

In contrast to the existing literature on firm-sponsored training, we are able to make use of detailed firm-level data on both hiring and training costs. This allows us to estimate simultaneously the effects of training costs and hiring costs on the firm's supply of apprentices.

3 Model

The model of the labor market used in this paper is an extended version of the generalized model of monopsony introduced by Manning (2006). In contrast to other models used in the literature on employment adjustment costs, the wage is no longer treated as exogenously given to the firm. $H(R, N, w)$ denotes the costs of recruiting and training a worker. First of all, these costs depend on the number of recruits R. At this point, we do not specify a functional form of H with respect to R. Instead, we allow the marginal hiring costs with respect to R to be increasing, constant or decreasing. As well, H depends on the number of skilled workers N that are already employed by the firm. On the one hand, large firms might have a more sophisticated recruitment process. On the other hand, firms with a large N might be more attractive because they provide better career opportunities, which would in turn make it easier for such firms to hire workers. The number of skilled workers N can be increased by either recruiting skilled workers R on the

labor market or by employing trainees L after their training period.
The wage w has two effects on H. Firms offering high wages are more attractive, hence more workers will apply to vacancies. However, a higher w also makes hiring more costly, since a worker does not reach his full level of productivity in the initial period after he has been hired. During this adaption period, a firm has to pay w, which is higher than the worker's productivity.
The costs of a trainee are denoted by $C(L)$. These costs are defined as net of the trainee's output contribution and independent of the number of skilled workers. Furthermore, we assume that the costs of training outweigh the trainee's output contribution during the training period. Therefore, skilled labor is the only production factor in our model.
The firm maximizes the present discounted value of its profits

$$\max_{R_t,w_t,L_t} \Pi = \sum_{t=0}^{\infty} \beta^t \left[F(N_t) - w_t N_t - H(R_t, N_t, w_t) R_t - C(L_t) L_t\right]$$

subject to the constraint

$$N_{t+1} = (1 - \delta(w_t))N_t + R_t + (1 - \gamma(w_t))L_t$$

where $F(N_t)$ denotes the firm's revenue function, $\delta(w)$ is the separation rate, i.e. the percentage of skilled workers that leaves the firm per period, with $0 \leq \delta(w) \leq 1$.[1] $\gamma(w)$ denotes the fraction of trainees which leave the firm after having completed their training, with $0 \leq \gamma(w) \leq 1$. The remainder, $(1 - \gamma(w))$, turns into skilled labor and stays in the firm.

We solve the problem of the firm by applying dynamic programming. Hence, we define the value function $V(N)$ to be the discounted value of profits if the employer has employment equal to N. The maximization problem in the Bellman form can then be written as:

[1] The separation rate is assumed to be continuous in the wage. We can in fact test this assumption and find an elasticity of -0.4, i.e., a 10% increase in the wage leads to a 4% decrease in the separation rate. This indicates that the firm can reduce the separation rate by paying a higher wage.

$$V(N_t) = \max_{R_t, w_t, L_t} F(N_t) - w_t N_t - H(R_t, N_t, w_t)R_t - C(L_t)L_t \\ +\beta V(N_{t+1}) \tag{1}$$

subject to the constraint

$$N_{t+1} = (1 - \delta(w_t))N_t + R_t + (1 - \gamma(w_t))L_t \tag{2}$$

where $\beta = \frac{1}{1+r}$ is the discount factor. Substituting (2) into (1) gives

$$V(N_t) = \max_{R_t, w_t, L_t} F(N_t) - w_t N_t - H(R_t, N_t, w_t)R_t - C(L_t)L_t \\ +\beta V((1 - \delta(w_t))N_t + R_t + (1 - \gamma(w_t))L_t)$$

Taking the first-order condition with respect to R_t yields

$$-H - R_t H_{R_t} + \beta V'(N_{t+1}) = 0 \tag{3}$$

The first-order condition with respect to w_t can be written as

$$-N_t - H_{w_t} R_t - \beta V'(N_{t+1})(\delta'(w_t)N_t + \gamma'(w_t)L_t) = 0 \tag{4}$$

The first-order condition with respect to L_t can be written as

$$-C_{L_t} L_t - C + \beta V'(N_{t+1})(1 - \gamma(w_t)) = 0 \tag{5}$$

To get a further optimality condition it is common to apply the Envelope Theorem in the Bellman context. In our case this is the derivative of the value function with respect to N_t:

$$V'(N_t) = F'(N_t) - w_t - H_{N_t} R_t + \beta V'(N_{t+1})(1 - \delta(w_t)) \tag{6}$$

In the steady state wages and employment are constant. This implies that we can rewrite (6) and solve for $V'(N)$:

$$V'(N) = F'(N) - w - H_N R + \beta V'(N)(1 - \delta(w)) \tag{7}$$

$$\Rightarrow \quad V'(N) = \left(\frac{1+r}{r + \delta(w)}\right)\left(F'(N) - w - H_N R\right) \tag{8}$$

Using steady state terms and the Envelope Theorem, condition (3) yields

$$H + RH_R = \beta V'(N) \tag{9}$$

$$\Rightarrow \quad H + RH_R = \left(\frac{1}{r+\delta(w)}\right)\big(F'(N) - w - H_N R\big) \tag{10}$$

Using steady state terms and the Envelope Theorem, condition (4) yields

$$-N - H_w R = \beta\left(\frac{1+r}{r+\delta(w)}\right)\big(F'(N) - w - H_N R\big)\left(\delta'(w)N + \gamma'(w)L\right) \tag{11}$$

which combined with (10) yields

$$-N - H_w R = (H + RH_R)(\delta'(w)N + \gamma'(w)L) \tag{12}$$

Using steady state terms and the Envelope Theorem, condition (5) yields

$$C_L L + C(L) = \beta\left(\frac{1+r}{r+\delta(w)}\right)\big(F'(N) - w - H_N R\big)\left(1 - \gamma(w)\right) \tag{13}$$

which combined with (10) yields

$$C_L L + C(L) = (H + RH_R)(1 - \gamma(w)) \tag{14}$$

Rearranging terms, (12) yields

$$H + RH_R = \frac{-(N + H_w R)}{(\delta'(w)N + \gamma'(w)L)} \tag{15}$$

which we combine with (14) to get

$$C_L L + C(L) = \frac{-(N + H_w R)}{(\delta'(w)N + \gamma'(w)L)}(1 - \gamma(w)) \tag{16}$$

So far we used a very general framework. Next, we assume more specific functions to learn more about the comparative statics in this model's steady state.

$$C(L_t) = c_t L_t \tag{17}$$

$$H(R_t, w_t, N_t) = w^\alpha R^\beta N^\gamma \tag{18}$$

$$\delta(w_t) = 1 - \xi w_t \tag{19}$$

$$\gamma(w_t) = 1 - \epsilon w_t \tag{20}$$

We assume a linear form of the training cost function C. The iso-elastic functional form of hiring costs is proposed by Manning (2006). The coefficients α, β and γ are exogenously given to an individual firm. The functional forms of the separation rates reflect the negative relation between the wage and the separation rates. Higher wages make it less attractive for workers to switch jobs.

With this functions, equation (16) can be written as

$$cL + cL = \frac{-(N + \alpha w^{\alpha-1} R^{\beta} N^{\gamma} R)}{(-\xi N - \epsilon L)} \epsilon w \tag{21}$$

Rearranging terms, we can rewrite this equation to get a quadratic solution for L

$$\begin{aligned} & -(\xi N + \epsilon L) 2cL = -(N + \alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}) \epsilon w \\ \Rightarrow \quad & (2\xi N c L + 2\epsilon c L^2) = (N + \alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}) \epsilon w \\ \Rightarrow \quad & 2\epsilon c L^2 + 2\xi N c L - (N + \alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}) \epsilon w = 0 \end{aligned}$$

Solving this quadratic equation yields

$$L_{1,2} = \frac{-2\xi N c \pm \sqrt{4\xi^2 N^2 c^2 + (8\epsilon c (N + \alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}) \epsilon w)}}{4\epsilon c}$$

Since we can not have a negative supply of training by a firm, there is obviously only one solution which makes sense:

$$L = \frac{-2\xi N c + \sqrt{4\xi^2 N^2 c^2 + (8\epsilon^2 c (N + \alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}) w)}}{4\epsilon c} \tag{22}$$

To get a non-negative expression for L, we infer that

$$\sqrt{4\xi^2 N^2 c^2 + (8\epsilon^2 c (N + \alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}) w)} \geq 2\xi N c$$

This is the case if

$$\left(8\epsilon^2 c (N + \alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}) w\right) \geq 0$$

or if

$$N \geq -\alpha w^{\alpha-1} R^{\beta+1} N^{\gamma}$$

This condition can be rewritten as

$$wN \geq -\alpha RH$$

If α is positive, this condition obviously holds. The condition is also fulfilled for negative values of α, if α is not unreasonably large in absolute values or if total wage costs wN are substantially higher than total hiring costs RH in a given period.

To sum up, we derived the firm's supply function of training positions, which is given by equation (22). Since we are interested in the effect of hiring costs H on the number of trainees L, we rewrite equation (22) as

$$L = \frac{-2\xi Nc + \sqrt{4\xi^2 N^2 c^2 + 8\epsilon^2 c(wN + \alpha RH)}}{4\epsilon c} \quad (23)$$

This expression indicates a positive relation between hiring costs H and the firm's supply of training positions L. This means that firms facing high hiring costs decide to offer training positions in order to cover part of their demand for skilled workers, instead of only recruiting skilled workers on the external labor market.

4 The apprenticeship system

The apprenticeship system has a long tradition in German speaking countries. It is characterized by the so-called dual education system, i.e., a training program combining training and working within a firm and vocational education at school. In Switzerland, this path is chosen by around 60% of young people who complete their compulsory schooling. From the remaining 40%, about one half attend grammar school to prepare them for an academic education whereas the remainder opt either for other entirely school-based forms of education or pursue no form of post-compulsory education. Apprentices can choose from over 200 different professions. Usually, an apprenticeship training program lasts three to four years. During this program, an apprentice spends about one or two days per week in a public vocational school. During the rest of the week, an apprentice participates in the production process or receives further training within the firm. After having completed the training program, apprentices receive a diploma recognized throughout the country. Graduated apprentices who have acquired an additional qualification (professional baccalaureate), have access to third-level studies at a university of applied sciences. Hence, an individual with a completed apprenticeship has various perspectives for further professional development.

In 2004, Swiss firms offering apprenticeships invested 4.7 billion Swiss francs in their training programs, which corresponds to about 1% of GDP. The apprentices generate a value of 5.2 billion Swiss francs during the training program (Muehlemann et al. 2007a). Hence, an apprenticeship training program is profitable for the firm on average. However, approximately one third of the apprenticeships end with net costs for the firm. If the employer and the apprentice want to continue their employment relationship after the training program, they have to negotiate a new labor contract. In Switzerland, only 37% of apprentices remain within the firm where they received their training one year after graduation. By continuing the employment relation with the graduated apprentices, the firm can reduce its recruitment costs for skilled labor. This way, firms may offset some training costs against the costs of hiring skilled workers on the external labor market.

5 Data

5.1 Survey design and data

The data used here are from two representative firm-level surveys conducted in Swiss firms in the years 2000 and 2004 by the Centre for Research in Economics of Education at the University of Berne and the Swiss Federal Statistical Office. Overall, we have information on hiring costs for a total of 4072 firms.[2]

The firms were asked about the number of skilled workers with a vocational degree that they have hired in the previous three years. The questionnaires were filled out either by management or the human resources department.

[2]Public firms and non-profit organizations have been excluded from the sample, since the principle of profit-maximization does not fully apply to those firms.

The answers reflect average costs of hiring a skilled worker with a vocational degree on the external labor market. The survey is a stratified random sample at the establishment level, where the two-digit industry level and the firm size serve as strata. The firms were asked to fill out hiring costs for a specific profession, which makes it easier to compare hiring costs across firms, since the comparisons can be made within a homogenous profession rather than across different occupations only.[3]

5.2 Calculation of hiring costs

The calculation of hiring costs can be divided into two parts: the costs of recruiting a worker, subsequently denoted by r, and the costs of initial training that is necessary to adapt to the new job, subsequently denoted by a.

Recruiting costs r_i are given by the costs for posting a vacancy v_i and interview costs $J_i c_{ai}$, where J_i denotes the number of applicants per vacancy that are invited for an interview and c_{ai} denotes the costs to conduct one interview, which are the product of interview time and the corresponding wage of the interviewer.[4] In addition, there are costs for external advisors or placement agencies e_i. Summarizing, recruitment costs can be written as

$$r_i = v_i + J_i c_{ai} + e_i$$

The second part of hiring costs arises because a newly employed skilled worker will not reach the full productivity from the beginning. In the questionnaire, firms were asked for how many days d_{ai} a new worker is less productive than an average skilled worker in the firm. The relative productivity is denoted by p_i. There are several reasons why a newly employed worker is less productive at the beginning. One explanation is firm specific human capital, which first has to be accumulated before a worker can reach full productivity. This may include factors such as getting to know the

[3] Firms were randomly assigned to fill out the questionnaire for a certain profession.

[4] There are five different job categories for interviewers: management, skilled workers with a vocational degree (by subcategories: administration, technical or social, crafts) and workers with no vocational degree.

firm culture, production processes and colleagues. Other reasons for lower productivity might be that newly hired workers receive training outside the workplace. This causes two types of costs for the firm: first, the firm has to pay the worker the daily salary w_{di} during the number of training days d_{ti}, and second, there are direct training costs c_{ti} for internal and external training personnel, travel costs or course fees. Hence, adaption costs a_i can be written as

$$a_i = d_{ai}(1 - p_i)w_{di} + d_{ti}w_{di} + c_{ti}$$

Finally, firm i's hiring costs[5] to fill a vacancy are given by

$$H_i = r_i + a_i$$

5.3 Calculation of the net costs of training apprentices

The net costs of apprenticeship training are given by the difference between the costs and benefits for the firm. The observed costs were calculated as follows: training costs are the wages of apprentices w_a and the cost for the training personnel w_T, which add up in equal shares to about 90 percent of total costs c. The remaining 10 percent are costs for material, infrastructure, external courses, hiring and administration of apprentices and other, x. This yields the following average costs for firm i:

$$c_i = w_{ai} + w_{Ti} + x_i$$

The calculation of training costs suggests that the main part of it is given by wages. Hence, training cost differences between firms are primarily due to variables that influence the wage level of either apprentices or training personnel. The wages of apprentices are more or less predetermined by wage recommendations of professional associations. Thus, the variation of

[5]Sometimes constructing a left-hand-side variable is criticized, but the only alternative would be to directly ask the firms about the monetary costs of hiring skilled workers. The problem with this approach is that firms might use different accounting procedures to internally calculate their costs. In our case, hiring costs are calculated in exactly the same way for all firms, wich makes comparisons much more meaningful.

wages within a profession is relatively small, and a substantial part can be explained by a firm's size. Larger firms offer higher wages to apprentices, which is consistent with the fact that larger firms offer higher wages to all categories of workers. Similarly to standard wage regressions, the average wages of the trainers can be best explained using variables such as firm size, industry and regional characteristics. The remaining costs for material, infrastructure and external courses are essentially given by the training profession.

The benefits b are calculated by the type of work the apprentices perform. An apprentice spends a fraction α of his work hours performing activities that would otherwise be done by unskilled workers. The remaining time $(1-\alpha)$ the apprentice carries out activities of a skilled worker. While we can assume in the first case that the apprentice's performance has the same value as that of an unskilled worker (w_u), the value of the apprentice's performance for the second case, ςw, is compared to that of a fully trained skilled worker, where w is the wage paid to a skilled worker.

$$b_i = \alpha w_{ui} + (1-\alpha)\varsigma w_i$$

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, as well as the profession in which an apprentice is trained. The net costs of training an apprentice, C, are the difference between the costs c and the benefits b:

$$C_i = c_i - b_i$$

5.4 Sample Selection

Firms can adjust their labor force N both by recruiting skilled workers R on the external labor market and by training apprentices L, which can be retained as skilled workers after the training period. Figure 1 illustrates the firms' possibilities to adjust N. Firms that did not hire externally and did not offer training ($L = 0$ and $R = 0$) have been excluded from the sample because they do not provide any relevant information.

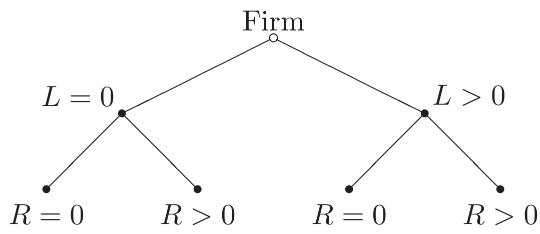

Figure 1: Hiring and training structure of the firm

The remaining cases have been included in the sample, but each case faces a sample selection problem. For firms with $L = 0$ and $R > 0$ we can observe H but not the costs for training C. For the case where $L > 0$ and $R = 0$, we cannot observe hiring costs H since no workers have been recruited externally. Finally, for firms where $L > 0$ and $R > 0$ we observe both H and C. However, the problem is that H does not refer to the same hiring costs as in the previous case, since firms that offer training may do so in order to reduce hiring costs for skilled workers. Hence, the variable we need is the potential hiring costs that would arise if a firm had to adjust its labor force solely on the external labor market. As a result, the observed H for firms with both $L > 0, R > 0$ is biased downwards.
Given the sample characteristics outlined above, we only use the observed values for H for firms where $L = 0$ and $R > 0$, whereas C is only observed for firms that offer training, i.e. $L > 0$.

5.5 Descriptive Statistics

Table A1 in the appendix reports the descriptive statistics. The number of apprentices L hired by firms is on average 0.9. The reason for such a low average supply of training positions is that only about one third of the firms train apprentices. But there is considerable variation in L. The highest number of apprentices trained by a firm is 134. The hiring costs H to fill a vacancy are between CHF 320 and CHF 17,0575 with a mean of CHF

14,285.[6] The net costs of training apprentices C are on average CHF -8,119, which means that training is profitable from the firms' perspective already at the end of the training period. However, about one third of the firms have to bear net costs of training. The separation rate of apprentices γ, i.e., the fraction of trained apprentices that leave the firm within the first year after completion of training, is on average 64%. The number of recruited skilled workers over the last three years R within a profession is 2.8 on average.[7] The firms' overall demand for skilled workers P is 3.2 on average, which is higher than R because it includes the apprentices that a firm retains as skilled workers after completion of training. The monthly salary of skilled workers is on average CHF 6,423.[8] The number of skilled workers N within the profession of interest is 6.7 on average, whereas the mean of employees other than N is 14.3.

6 Econometric model and empirical analysis

In this section we estimate the effect of hiring costs on the supply of training positions. Since firms supplying training may have reduced recruitment costs for skilled labor, we cannot observe their potential hiring costs. Hence, we need to estimate these unobservable costs. In order to estimate the effect of hiring costs on the firm's supply of training, we need to control for the net costs of training. However, these costs can only be observed for firms that offer training. The expected net costs for non-training firms have to be estimated.

To overcome the selection problems mentioned above, we use a type IV Tobit model, which is described in the next subsection.[9]

[6]Detailed descriptive statistics of the components of H can be found in Table A2.

[7]Firms that train apprentices had to report training costs and hiring costs for the same profession. Firms that do not train apprentices were asked to report hiring costs for the profession in which they would train apprentices if they decided to do so.

[8]It should be noted that w in the theoretical model refers to the wage over a period of three years.

[9]For a classification of different types of Tobit models see Amemiya (1985).

6.1 Type IV Tobit model

Consider the (structural) model of the firm's supply of training positions.

$$\ln H = x_1\beta_1 + \varepsilon_1 \tag{24}$$

$$C = x_2\beta_2 + \varepsilon_2 \tag{25}$$

$$L = \max[0, x_3\beta_3 + \alpha \ln H + \delta C + \varepsilon_3] \tag{26}$$

where again L denotes the number of trainees, H the costs of hiring skilled workers and C the net costs of training.

(x, L) is always observed; H is observed if $L = 0$, whereas C is observed if $L > 0$. $(\varepsilon_1, \varepsilon_2, \varepsilon_3)$ are independent of x with a zero-mean trivariate normal distribution. Furthermore, x_1 needs to contain at least one element that is not in x_3 and has a non-zero coefficient. In this case we use the binary variable *Difficulties to find skilled workers*. This variable is a measure of the tightness of the labor market.[10] If skilled labor is scarce, firms have to spend more resources on finding appropriate workers. This increases the hiring costs of the firm. However, we assume that a firm's difficulties to find skilled workers has no direct impact on the supply of training positions. Similarly, x_2 contains at least one element that is not in x_3 and has a non-zero coefficient. In this case we use the variable *Local share of young people*. This variable measures the share of the 15-19 years old people of the total population in a region.[11] The share of young people influences the net costs of training for the following reason. If there are many potential trainees for a firm, the probability of a firm to find a good match increases. This in turn reduces the firm's time and effort associated with training. However, we assume that the share of young people does not directly affect the firm's supply of training positions.
The coefficients of interest are the structural parameters α and δ. For identification, we follow the procedure described in Wooldridge 2002, pp.571.

[10] See Majumdar (2007) for an analysis of the effect of labor market conditions on the firm's incentive to offer training.

[11] A region is defined to contain all cities that can be reached by car within 30 minutes from a certain city.

The reduced form for equation (26) is

$$\begin{aligned} L &= \max[0, x_3\beta_3 + \alpha(x_1\beta_1 + \varepsilon_1) + \delta(x_2\beta_2 + \varepsilon_2) + \varepsilon_3] \\ &= \max[0, x\rho_3 + \alpha\varepsilon_1 + \delta\varepsilon_2 + \varepsilon_3] \\ &= \max[0, x\rho_3 + u_3] \end{aligned} \tag{27}$$

In a first step we regress L on x by standard Tobit using all observations. This enables us to generate a generalized residual (see Vella 1992, p.418):

$$\hat{u}_{3i} = -\hat{\sigma}_3(1 - I_i)\phi(x_{3i}\hat{\beta}_3/\hat{\sigma}_3)(1 - \Phi(x_{3i}\hat{\beta}_3/\hat{\sigma}_3))^{-1} + I_i(L_i - x_{3i}\hat{\beta}_3)$$

where I_i is an indicator function denoting whether a firm offers training. A consistent estimate of u_{3i} is necessary to obtain estimates of β_1 and β_2, since

$$E(\ln H|L = 0, x, u_3) = x_1'\beta_1 + E(\varepsilon_1|x, u_3) = x_1'\beta_1 + E(\varepsilon_1|u_3) = x_1'\beta_1 + \gamma_1 u_3$$

and

$$E(C|L > 0, x, u_3) = x_2'\beta_2 + E(\varepsilon_2|x, u_3) = x_2'\beta_2 + E(\varepsilon_2|u_3) = x_2'\beta_2 + \gamma_2 u_3$$

where $u_3 = \alpha\varepsilon_1 + \delta\varepsilon_2 + \varepsilon_3$. The coefficients γ_1 and γ_2 can be used to test for selectivity (see also Vella 1992).

Using observations for which $L > 0$, we estimate the OLS regression

$$\ln H_i \quad \text{on} \quad x_{i1}, \hat{u}_{3i} \tag{28}$$

This yields consistent estimates of β_1 and allows to test for selectivity bias. Similarly, using observations for which $L = 0$, we estimate the OLS regression

$$C_i \quad \text{on} \quad x_{i2}, \hat{u}_{3i} \tag{29}$$

This yields consistent estimates of β_2 and again allows to test for selectivity bias.

Now we have consistent estimates $\hat{\beta}_1$, $\hat{\beta}_2$ for β_1, β_2. This enables us to estimate β_3, α and δ using the following reduced form of L in terms of the structural parameters:

$$L = \max[0, x_3\beta_3 + \alpha\ln(x_1\beta_1) + \delta(x_2\beta_2) + u_3]$$

Using our consistent estimates $\hat{\beta}_1$, $\hat{\beta}_2$ we can estimate the following Tobit equation to obtain consistent estimates $\hat{\beta}_3$, $\hat{\alpha}$ and $\hat{\delta}$, which is what we are interested in:

$$L = \max[0, x_3\beta_3 + \alpha \ln(x_1\hat{\beta}_1) + \delta(x_2\hat{\beta}_2) + error_i] \qquad (30)$$

It should be noted that the explanatory variables in equation (30) are $x_1\hat{\beta}_1$ and $x_2\hat{\beta}_2$ for all i and do not depend on $\hat{u}_{3i}$ in equations (28) and (29).

6.2 Results

First, we calculate the generalized residual by estimating equation (27).[12] Second, we estimate equation (28), i.e., we regress hiring costs $\ln H$ on x_1 and the generalized residual. The results are reported in Table A4. Using only observations for which the potential hiring costs are observable (and therefore $L = 0$), we find that the firm's total demand for skilled workers $\ln P = \ln[R + (1-\gamma)L]$[13] has a positive and significant effect on hiring costs with an elasticity of 0.161.

As an exclusion restriction, i.e., a variable that has an effect on H but not directly on L, we included a dummy variable measuring the difficulties of a particular firm to find skilled workers on the external labor market. As expected, firms facing such difficulties exhibit about 23% higher hiring costs. The coefficient on the generalized residual is not significantly different from zero. Hence, there does not seem to be a selection problem, i.e. the potential hiring costs of firms with $L > 0$ do not differ significantly from firms with $L = 0$.[14] Furthermore, the firm size has a positive effect on H, although the number of skilled workers N is not significant alone.[15]

Next, we estimate equation (29), i.e., we regress the net costs of training C on x_2 and the generalized residual.

[12] The results of the reduced form equation are presented in Table A3.

[13] Note that in this case $P = R$, since $L = 0$, i.e., the firm does not offer training.

[14] However, for reasons of consistency we include the generalized residual because there is a selection bias with regards to the net costs of training C (see Table A5).

[15] A Wald test for joint significance for total firm size (N + employees other than N) yields a p-value of 0.02.

The estimation results are presented in Table A5. The exclusion restriction used here is the share of young people on the population in a region. The results show that a 1%-point increase in the share of young people leads to a decrease in the net costs of training of CHF 3761. The intuition for this result is that firms are more likely to find a good match if the supply of young people is high. Therefore, they can save on training costs compared to firms that cannot find suitable trainees.

The coefficient on the generalized residual is now negative and significantly different from zero. This leads to the conclusion that the expected net costs of training for firms with $L = 0$ are significantly higher than those for firms with $L > 0$, indicating a sample selection problem. This confirms earlier findings of Wolter et al. (2006) and Muehlemann et al. (2007). In addition, the firm size has a positive effect on C. The number of skilled workers N and employees other than N are jointly significant at the 1%-level.

As a final step, we estimate the structural effects of H and C on the firm's supply of training positions L. The results show that both variables have a significant effect (see Table 1). Hiring costs of skilled workers have a positive effect on L. If H increases by CHF 1000, L increases by 0.25, implying that an increase in average hiring costs of CHF 4,000 induces a firm to hire an additional trainee. This effect is economically substantial, since an increase of H by one standard deviation leads to an increase of 3.5 training positions offered by a firm.

As expected, the net costs of training have a negative effect on the firm's supply of training. If the net costs C increase by CHF 1000, L decreases by 0.18. Similarly, an increase in C by one standard deviation leads to a decrease of 7 training positions on average.

Furthermore, the number of skilled workers and the number of other employees within the firm both have a positive effect on L. The wage of skilled workers w has a negative effect on L. A possible explanation is that the wage is negatively related to the separation rate of trainees. Hence, a firms needs to train less apprentices in order to fill a given number of vacancies. Finally, foreign-owned firms have a significantly lower supply of training. This confirms the results of earlier studies. A possible explanation for this

Table 1: Supply of training positions

Dependent variable: L	
$\hat{H}$ (in 1000)	0.2497
	(0.0323)
$\hat{C}$ (in 1000)	-0.1837
	(0.0636)
P	-0.0188
	(0.0158)
N	0.0410
	(0.0064)
Employees other than N	0.0125
	(0.0022)
w	-0.0010
	(0.0002)
Foreign firm ownership	-1.1907
	(0.5231)
Aggregate cantonal income (in 1'000 CHF)	0.0064
	(0.0170)
Industry controls	Yes
Job controls	Yes
Log pseudolikelihood	-296125
Observations	4511

Cluster-robust standard errors in parentheses.

Table 2: Supply of training positions, $\hat{C} > 0$

Dependent variable: L	
$\hat{H}$ (in 1000)	0.4006
	(0.0730)
$\hat{C}$ (in 1000)	-0.2106
	(0.2216)
P	0.1604
	(0.0532)
N	0.0132
	(0.0078)
Employees other than N	0.0093
	(0.0039)
w	-0.0014
	(0.0004)
Foreign firm ownership	-1.2766
	(1.2605)
Aggregate cantonal income (in 1'000)	-0.0079
	(0.0457)
Industry controls	Yes
Job controls	Yes
Log pseudolikelihood	-67400
Observations	1331

Cluster-robust standard errors in parentheses.

result is that such firms might be less familiar with the vocational training system or too specialized to provide a complete training program.

We have also estimated the supply of training positions using only firms with positive net costs C. Since the firm's total demand for skilled workers P depends on L by construction, because $P = R + (1-\gamma)L$, and β_P is significantly different from zero, the coefficients in the regression output cannot be interpreted directly. Hence, we rewrite

$$\begin{aligned} L &= \beta_C C + \beta_H H + \beta_P P + \beta_X X + \beta_{[HC]} HC + u \\ &= \beta_C C + \beta_H H + \beta_P (R + (1-\gamma)L) + \beta_X X + \beta_{[HC]} HC + u \\ &\Leftrightarrow \\ L &= \frac{1}{1-\beta_P(1-\gamma)} \big(\beta_C C + \beta_H H + \beta_P R + \beta_X X + \beta_{[HC]} HC + u\big) \end{aligned}$$

Using this transformation, we find that an increase in hiring costs H of CHF 1000 yields an increase in the supply of training positions L of 0.425. Put differently, an increase in H by one standard deviation leads to an increase of 6 training positions supplied by the firm. The effect of H is stronger if we only consider firms with $C > 0$, because hiring costs are not the deciding factor if the firm does not need to recover any training investments after the training period.

Summarizing, the empirical results show that firms facing high hiring costs decide to offer more training in order to satisfy their demand for skilled workers compared to firms with low hiring costs.

7 Conclusions

Firm-sponsored training provides an alternative way to recruit skilled workers and allows firms to avoid high hiring costs. Our empirical results provide evidence that the firm's supply of training in fact depends on the firm's level of hiring costs. We find that an increase in average hiring costs has a substantial positive impact on the firm's supply of training positions. An increase in hiring costs by one standard deviation leads a firm to offer 6 additional training positions. This is an explanation for the well-known observation that firms are frequently willing to bear substantial training costs. Especially in times of shortages in skilled labor, when hiring costs are particularly high, it is beneficial for firms to supply training positions and retain their former trainees as skilled workers after training. A possible implication of our results is that firms are more willing to offer training positions if labor markets are strongly regulated. For example, a high degree of employment protection legislation forces firms to invest more in their hiring activities in order to avoid costly mismatches. Hence, training young workers internally may become an appropriate strategy for a firm to satisfy its demand for skilled workers.

References

Abowd, John M. and Francis Kramarz (2003). The costs of hiring and separations. *Labour Economics 10*(5), 499–530.

Amemiya, Takeshi (1985). *Advanced Econometrics*. Harvard University Press, Cambridge, Massachusetts.

Kramarz, Francis and Marie-Laure Michaud (2004). The Shape of Hiring and Separation Costs. IZA Discussion Paper No.1170.

Majumdar, Sumon (2007). Market conditions and worker training: How does it affect and whom? *Labour Economics 14*, 1–23.

Manning, Alan (2006). A Generalised Model of Monopsony. *Economic Journal 116*(508), 84–100.

Muehlemann, Samuel, Juerg Schweri, Rainer Winkelmann, and Stefan C. Wolter (2007). An Empirical Analysis of the Decision to Train Apprentices. *Labour 21*(3), 419–441.

Stevens, Margaret (1994). An Investment Model for the Supply of Training by Employers. *Economic Journal 104*(424), 556–70.

Vella, Francis (1992). Simple Tests for Sample Selection Bias in Censored and Discrete Choice Models. *Journal of Applied Econometrics 7*(4), 413–21.

Wolter, Stefan C., Samuel Muehlemann, and Juerg Schweri (2006). Why Some Firms Train Apprentices and Many Others Do Not. *German Economic Review 7*(3), 249–264.

Wooldridge, Jeffrey M. (2002). *Econometric analysis of cross section and panel data*. MIT Press.

A Tables

Table A1: Descriptive Statistics

	Mean	Std. Dev.	Min	Max	Obs
Number of trainees L	0.904	2.351	0	134	4511
Hiring costs H	14284.710	14560.850	320	170575	1675
Net costs of training C	-8118.541	35048.480	-132707	149155	2836
Separation rate of trainees γ	63.771	34.470	0	100	2836
Number of recruits R	2.815	8.283	0	450	4511
Demand for skilled workers P	3.245	7.748	1	461	4511
Wage of skilled workers (per month)	6423.296	1450.822	2300	14430	4511
Number of skilled workers N	6.743	24.153	1	2400	4511
Number of employees other than N	14.318	74.916	0	4610	4511
Foreign-owned firm	0.120		0	1	4511
Construction sector	0.128		0	1	4511
Industrial sector	0.134		0	1	4511
Service sector	0.738		0	1	4511
Aggregate cantonal income	48949.400	10310.320	33699	82415	4511
Professional dummies:					
Administrative assistant	0.226		0	1	4511
Electrician	0.028		0	1	4511
Polymechanics technician	0.022		0	1	4511
Sales clerk	0.044		0	1	4511
Cook	0.061		0	1	4511
Local share of young people	0.057	0.005	0.047	0.071	4511
Difficulties to find skilled workers	0.395		0	1	4511
Year of survey (1=2000, 0=2004)	0.481		0	1	4511

Table A2: Descriptive statistics

Variable	Mean	Std.Err.	Minimum	Maximum	Obs.
Costs for job postings v (in CHF)	1103	1889	0	50000	4032
Costs for interview per applicant c_a (in CHF)	395	495	0	8844	4032
Number of interviewed applicants J per vacancy	5	4	1	30	4032
Personnel costs for interviews $J * c_a$	2009	3877	0	83586	4032
Costs for external advisors/headhunters e (in CHF)	414	1881	0	30000	4032
Recruitment costs $r = v + J * c_a + e$ (in CHF)	3878	5894	0	116117	4032
Duration of adaption period in days d_a	80	60	0	756	4032
Average decline in productivity $(1 - p)$ during adaption period (in %)	29	14	0	90	4032
Daily wage w of a skilled worker with vocational degree (in CHF)	349	79	125	784	4032
Duration of training courses in days d_t	2	4	0	90	4032
Direct training costs c_t (in CHF)	550	1805	0	60000	4032
Adaption costs $a = d_a * (1 - p)w + d_t * w + c_t$ (in CHF)	9688	11005	0	147779	4032
Average hiring costs $H = r + a$ to fill a vacancy (in CHF)	13570	13862	320	170575	4032

Table A3: Reduced form Tobit

Dependent variable: L	
P	0.0247
	(0.0169)
N	0.0432
	(0.0069)
Employees other than N	0.0112
	(0.0019)
w	-0.0001
	(0.0001)
Foreign firm ownership	-1.8638
	(0.2874)
Aggregate cantonal income (in 1'000 CHF)	-0.0073
	(0.0101)
Local share of young workers	69.4752
	(18.5540)
Difficulties to find skilled workers	0.8624
	(0.1771)
Industry controls	Yes
Job controls	Yes
Log pseudolikelihood	-298284
Observations	4511

Cluster-robust standard errors in parentheses.

Table A4: Hiring cost regression

Dependent variable: $\ln H$	
$\ln P$	0.1611
	(0.0715)
$\ln N$	0.0166
	(0.0529)
ln(Employees other than N)	0.0209
	(0.0067)
$\ln w$	1.4741
	(0.1214)
Foreign firm ownership	0.0129
	(0.0795)
Aggregate cantonal income (in 1'000 CHF)	0.0052
	(0.0024)
Difficulties to find skilled workers	0.2332
	(0.0565)
Generalized residual	0.0098
	(0.0293)
Industry controls	Yes
Job controls	Yes
R^2	0.2676
Observations	1675

Cluster-robust standard errors in parentheses.

Table A5: Net cost regression

Dependent variable: C	
P	-123.4067
	(41.8482)
N	22.4124
	(15.5346)
Employees other than N	17.7061
	(5.3174)
Foreign firm ownership	4755.4220
	(3831.6810)
Aggregate cantonal income (in 1'000 CHF)	171.9899
	(87.4384)
Local share of young workers (in %)	-3761.7190
	(1551.0780)
Generalized residual	-824.8130
	(257.2791)
Industry controls	Yes
Job controls	Yes
R^2	0.0886
Observations	2836

Cluster-robust standard errors in parentheses.