AN ALTERNATIVE APPROACH TO ESTIMATE THE WAGE RETURNS TO PRIVATE-SECTOR TRAINING

EDWIN LEUVEN[a,b]* AND HESSEL OOSTERBEEK[a,b]

[a] *Amsterdam School of Economics, University of Amsterdam, Amsterdam, The Netherlands*
[b] *Tinbergen Institute, Amsterdam, The Netherlands*

SUMMARY

This paper follows an alternative approach to identify the wage effects of private-sector training. The idea is to narrow down the comparison group by only taking into consideration the workers who wanted to participate in training but did not do so because of some random event. This makes the comparison group increasingly similar to the group of participants in terms of observed individual characteristics and the characteristics of (planned) training events. At the same time, the point estimate of the average return to training consistently drops from a large and significant return to a point estimate close to zero. Copyright © 2008 John Wiley & Sons, Ltd.

1. INTRODUCTION

The empirical literature on private-sector training addresses two main questions: who gets training, and what is it worth? While the first question has been answered fairly satisfactorily, this is not the case for the second question. Even when attention is restricted to the effect of training on wages rather than on productivity, severe problems are posed by the endogeneity of training decisions.[1]

Workers who participate in training, or who participate more often or for longer durations, are unlikely to have the same characteristics as other workers. As with regular education it seems likely that the selection of workers into training is correlated with workers' unobserved ability. If this is the case, and since ability positively affects wages, a regression of wages on training will not produce a causal effect but suffer from so-called ability bias. A key issue in the estimation of returns to training is thus how to correct for this potential source of bias.

The empirical literature on private-sector training contains basically two approaches. The first approach is to augment the wage equation with a Heckman-type selection correction term which results from a first-stage training participation equation. Results from this approach are reported by Lynch (1992) and Veum (1995), among others. The difficulty with this approach is twofold. First, the parametric selection models estimated in the literature are restrictive in the sense that they make an assumption about the distribution of the unobservables. Second, and more importantly, it is very hard to find variables which affect training participation and have arguably no direct effect on wages. The problem of finding such credible exclusion restrictions also hampers the application

---

* Correspondence to: Edwin Leuven, Amsterdam School of Economics, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands. E-mail: e.leuven@uva.nl

[1] To avoid confusion, the term 'private-sector training' refers to training courses undertaken by firms and their employees without any direct public intervention. A different literature studies public-sector training programs which are almost exclusively targeted at the unemployed, Heckman *et al.* (1999b) is an authoritative review of this literature.

of an instrumental variable (IV) approach. A second, more often used approach is to estimate the wage return to work-related training using fixed-effects regressions. This estimator, which is similar in spirit to taking first differences of the before-and-after training log wages, purges permanent individual effects from the estimating equation. Examples of studies that follow this approach include Barron *et al.* (1993), Booth (1993), Frazis and Loewenstein (2005), Greenhalgh and Stewart (1987), Lynch (1992), Parent (1999) and Veum (1995).

The fixed-effects estimator produces unbiased estimates when the unobserved individual effects are permanent. It is conceivable, however, that, apart from selection based on fixed individual observables and unobservables, selection into training also has dynamic aspects that provide an additional potential source of bias. Consider, for example, the case where individuals decide to take training because their earnings are temporarily low; faster earnings growth is then expected to occur among the trainees even in the absence of training participation. More generally, fixed-effects estimations do not recover causal relationships if wage growth is different for trainees and non-trainees. It should be noted that fixed-effect estimates of wage returns to training are typically smaller than standard OLS estimates, suggesting that fixed-effect estimates at least partially eliminate selection bias.

Estimates of the wage returns to private-sector training are typically quite high. As an illustration take the estimates from Frazis and Loewenstein (2005), who use the NLSY dataset and present a careful and thorough analysis of these data. They estimate various specifications, and their preferred estimate that takes into account heterogeneity in wage growth is a rate of return in the region of 40–50% for one full-time week of training.

The estimated wage return to a training spell of a median length of 60 hours equals 2.5%. Using the EOPP data, the authors find a wage return of 5% for a median training spell of 38 hours. These findings are consistent with those of Barron *et al.* (1993) and Loewenstein and Spletzer (1999a).[2]

Estimated returns are also high with data from other sources and countries. Bartel (1995) using company data, for example, finds that one day of training increases wages by 2%, which in her data is equivalent to a rate of return of 60%. Blundell *et al.* (1996) report returns to training incidence (zero-one dummy variable) for men in the UK in the region of 8% using OLS estimations, 9% for fixed-effects estimations, and 7% for IV estimations. The returns for women are even higher.

These results illustrate the fact that for a variety of datasets and countries the estimated returns to private-sector training are substantial. Moreover, the returns to private-sector training are very high compared to, for example, the returns to education. The return to a *year* of full-time education is around 10%, where in contrast the literature often finds returns of at least 3% for a *week* of private-sector training. This raises the question whether these estimates are indeed causal effects.[3]

In this paper we follow a different approach to estimate the returns to private-sector training. We use OLS as a benchmark result that does not correct for selectivity on unobservables. We then compare these results with estimates based on our approach that takes the concept of random assignment literally. The idea is to narrow down the comparison group to those non-participants who did not participate due to some random event. This is achieved by using the information obtained through two specially designed survey questions. The first is whether there was any

---

[2] Lynch (1992) and, to some extent, Parent (1999) find lower returns using the NLSY data. According to the analysis in Frazis and Loewenstein (2005) this is due to the linear specification of training, which leads to a serious underestimation of the return to training.

[3] There are, of course, exceptions. Some studies find smaller returns, typically for continental European countries (e.g., Pischke, 2001; Goux and Maurin, 2000), although other studies find larger returns for the same countries (e.g., Fougère *et al.*, 2001; A. Kuckulenz and T. Zwick, unpublished MS, 2003).

training related to work or career that the respondent wanted to attend but did not do so. The second asks whether this non-participation was due to some random event such as family circumstances, excess demand for training places, transient illness, or sudden absence of a colleague. Respondents who give an affirmative answer to both questions are arguably a more appropriate comparison group. If the random event acts as random assignment, this approach gives an estimate of the effect of treatment on the treated. One advantage of the approach followed in this paper is that it can be implemented using a single cross-section.

OLS gives us an estimate that is similar in magnitude to those found for the studies cited above, and is 9.5% for participating in one training course (with median duration of 40 hours) during the past 12 months. Restricting the comparison group to workers who wanted to participate in training but did not do so reduces the estimated return to 6.3%. When the comparison group is further restricted to those workers who wanted to participate in training but did not do so due to some random event, the point estimate of the wage return to training equals −1%, and is statistically insignificant. Unfortunately the small size of our preferred comparison group does not allow precise estimation of the latter effect, and we cannot exclude some of the more modest estimates in the literature. The credibility of the proposed strategy is, however, supported by the fact that narrowing down the comparison group to those who wanted to participate makes them more comparable to participants in terms of observed individual characteristics. Subsequently restricting the comparison group to those who did not participate due to some random event makes the comparison group also more similar in the characteristics of the (planned) training events.

We should stress that actual training courses can vary a lot in quality, duration and type, and that the effects of any particular training course will also vary with the characteristics of the employee getting the training. Like previous studies, our approach does not evaluate the effect of a specific training course for a specific worker, but the average effect of all courses on all participants. To put the findings into perspective it is informative to realize that the median duration of a training course in our sample (which is a representative sample of the Dutch working population aged 16–64) amounts to 40 hours, that most courses are financed by the employer, and that most courses lead to a certificate for the participants.

The next section discusses at greater length the questions that were used to create the new comparison group and how these help to estimate the wage effects of training courses. Section 3 presents the data and compares participant and comparison groups in terms of observed characteristics of employees and training courses. Section 4 presents the empirical results. Section 5 concludes.

2. METHOD

Studies that estimate causal wage returns to private sector training compare wages of employees that participated in training to the wages of an appropriate comparison group of employees. The training measure used in this paper to define the group of participants is a conventional one.[4] The exact phrasing of the question that is used to determine this reads:

> Did you spend time attending a course/training for purposes of your work or career opportunities during the past 12 months?

---

[4] To illustrate, the training question in the NLSY reads 'Since [the date of the last interview], did you attend any training program or any on-the-job training designed to help people find a job, improve skills, or learn a new job?'

All respondents who received private sector training during the 12 months prior to the interview are assigned to what we will refer to as Participant group I. Without any correction for selectivity the comparison group consists of all respondents who did not attend a course or training during the 12 months prior to the interview. We refer to this group as Comparison group I.

If training participation is randomly assigned to workers, the difference between the average wage in Participant group I and Comparison group I gives the causal effect of training on wages. It is unlikely, however, that training is assigned on a random basis. Selection into training requires that (1) the worker is willing to undertake training, and (2) the employer is prepared to provide it. The factors underlying these selection mechanisms are likely to be related, directly or indirectly, to potential outcomes and may lead to differences between the training participants and possible comparison groups in terms of characteristics that are not observed by the analyst. As a result, comparing wages of Participant group I and Comparison group I will give a biased estimate of the causal effect of training on wages.

The identification strategy that is proposed in this paper reduces the comparison group to those workers who are willing to undertake training and whose employers are prepared to provide it, but who did not attend the training course they wanted to attend due to some random event. Narrowing down the comparison group proceeds in two steps. The first step reduces the comparison group to the group of untrained workers who wanted to attend training but did not do so. This is done on the basis of information from a question which asks respondents the following:

> Was there any course/training related to work or career you wanted to attend but did not during the past 12 months?

Persons who respond that there was such a course, and who did not receive any training at all during the past 12 months, are assigned to Comparison group II. Note that Comparison group II is a subsample of Comparison group I. Comparison group II is arguably a more suitable comparison group than Comparison group I as it singles out all workers who were motivated to participate in training. Hence, it takes care of the first of the two selection mechanisms.

The second step is to further reduce the comparison group to untrained workers whose non-participation is due to some random event. Respondents who indicated their intention to be trained (i.e., wanted to attend some course/training but did not do so) were then asked the reason for not attending the course/training. To answer this question respondents had to choose one out of five alternatives:

1. A random event ($N = 77$)
2. Lack of time ($N = 93$)
3. Own financial contribution too high ($N = 13$)
4. Lack of support from the employer ($N = 21$)
5. Other reasons ($N = 45$)

Among all 249 respondents who indicated that there was a training course they would have wanted to receive, there are 77 respondents who say that they did not do so due to some random event. The respondents were given the following examples of such events: family circumstances, transient illness, or sudden absence of a colleague. These persons constitute the final comparison group, referred to as Comparison group III.

Comparison group III consists of respondents who did not attend any training course at all during the 12 months prior to the interview due to some random event. The phrasing in the questionnaire refers to a single training course that these respondents did not attend. Participant group I, however, consists of respondents who received at least one course. Comparison group III therefore seems a more appropriate comparison group for the group that received exactly one training/course than for the group who received two or more courses. For this reason Participant group II is constructed, which consists of respondents who received exactly one training/course.

Note that, of the three training measures—participation, number of courses and number of hours—the last one is the probably the most accurate measure of the investment in human capital. But since the questionnaire did not ask respondents the length of the training they missed because of some random event, it is not possible to estimate a causal effect at this margin. The analysis is therefore based on the other two measures. With Participant group I training is measured as mere participation, whereas with Participant group II training is measured as one course versus no training at all. Table I summarizes the definition of the participant and comparison groups.

Given that the assignment of respondents to Comparison group III is crucial for the approach of this paper, note that respondents are not assigned to the final comparison group when they mention one of the other categories as a reason for not having attended training. These other categories include the more obvious ones such as lack of time, too expensive and lack of employer support. They also include the category 'other reasons'. This is an open category which interviewees had to respond to when they mentioned 'other reasons'. In the category 'other reasons' the following reasons for not participating were mentioned: language problems, merger of current employer, no available transportation, change of job, moving house, stay abroad, pregnancy. The respondents who mentioned these reasons did not therefore consider these events as random.

Note further that in the absence of the random event which withheld them from training the respondents in Comparison group III would have participated in training. Comparison group III therefore serves to identify the effect of the treatment on the treated.

The approach followed here differs from the use of 'no-shows' as a comparison group as has been done for the evaluation of active labor market programs (Bell *et al.*, 1995; see also Heckman *et al.*, 1999a, p. 1940). No-shows are applicants to the program who have been accepted but nevertheless fail to participate in the program. Because the reasons for this non-participation are unknown, it may be related to systematic but unobserved characteristics which may thus lead to biased estimates. Translated to our application, workers belonging to Comparison group II with the exception of those who mention lack of support from their employer would constitute the group of no-shows. Going from Comparison group II to our preferred Comparison group III attempts to delete those cases from the comparison group for whom non-participation is likely to be related to non-random factors.

Table I. Definition of the participant and comparison groups

	Definition
Participant I	At least one training course
Participant II	Exactly one training course
Comparison I	No training
Comparison II	No training, but wanted to receive training course
Comparison III	No training, but wanted to receive training course and did not do so because of a random event

3. DATA

The data were collected in January and February 2001. Interviews were held by telephone using computer-aided techniques. The data are a representative sample of the Dutch population aged 16–64. The employed persons were asked questions concerning their employment characteristics. They also responded to an extensive set of questions about the training activities they undertook in the 12 months prior to the interview.

Table II presents sample means for the two participant groups and the three comparison groups. These means relate to gender, age, education, firm size, number of children, being non-Dutch, being single, temporary job status and firm tenure. These variables are often included in wage equations as controls. The empirical analysis in the next section presents results from wage equations with and without controls for these variables.

The dependent variable in the wage equations is the log of the hourly wage rate. Wages are measured as current gross wage income per month.[5] This is then divided by the number of (contractual) working hours per month. Because wages are measured as wages in the month of the interview and (potential) training participation is measured as training during the 12 months prior to the interview, this implies that on average the time that has elapsed between training completion and measurement of the wage rate will be in the neighborhood of 6 months. This relative timing of measurement of training and wages is similar to how this is done in most other empirical training studies (e.g., Frazis and Loewenstein, 2005; Loewenstein and Spletzer, 1999b; Parent, 1999). Interestingly, in one of the few studies that allows returns to vary with the amount of time passed since the moment of training, Lengermann (1999) finds that the largest part of the training returns materialize within 1 year of the training course. Like most studies he uses standard fixed-effects techniques.

Table II. Sample means per participant and comparison group

| | Participant | | Comparison | | |
	I (1)	II (2)	I (3)	II (4)	III (5)
Female	0.38	0.39	0.43	0.43	0.50
Age	36.25	35.64	37.98	36.54	36.52
Children	0.81	0.83	0.90	0.99	0.98
Non-Dutch	0.06	0.06	0.06	0.07	0.03
Single	0.16	0.15	0.14	0.16	0.20
Temporary job	0.18	0.22	0.19	0.18	0.20
Firm tenure (months)	102.51	100.97	109.70	89.16	89.65
Education					
Low	0.12	0.14	0.20	0.15	0.10
Intermediate	0.48	0.49	0.52	0.57	0.61
High	0.40	0.37	0.28	0.29	0.28
Firm size					
Up to 50	0.33	0.36	0.42	0.42	0.39
50–200	0.25	0.24	0.22	0.23	0.25
More than 200	0.42	0.40	0.36	0.34	0.36
N	1021	582	1145	249	77

---

[5] In the Netherlands people typically earn monthly salaries; only 4% of respondents report wages per week.

Table III. Tests of equality between participant (P) and comparison (C) groups, p-values

	PI vs. CI (1)	PI vs. CII (2)	PI vs. CIII (3)	PII vs. CI (4)	PII vs. CII (5)	PII vs. CIII (6)
Female	0.015	0.224	0.061	0.087	0.364	0.091
Age	0.002	0.730	0.844	0.000	0.319	0.528
Children	0.068	0.046	0.325	0.218	0.094	0.391
Non-Dutch	0.953	0.615	0.309	0.984	0.601	0.342
Single	0.371	0.965	0.491	0.640	0.815	0.431
Temporary job	0.693	0.981	0.712	0.429	0.381	0.785
Firm tenure	0.160	0.061	0.223	0.141	0.126	0.303
Education	0.000	0.010	0.078	0.000	0.110	0.144
Firm size	0.000	0.003	0.253	0.006	0.035	0.481

Note: The p-values are based on t-tests for the continuous variables age, number of children, firm tenure and log wage and on rank-sum tests for the categorical variables female, education, firm-size, non-Dutch, single and temporary job.

The means reported in Table II already hint at the fact that Comparison groups II and III are more comparable to Participants groups I and II than Comparison group I. By and large the means of Comparison groups II and III are closer to those of Participants groups I and II than the means of Comparison group I. This is most notably the case for the variables age, education and firm size. With the exception of the variables female and children, the means of the other variables in Table II are not very different across all five groups. Formal test statistics are reported in Table III.

Table III reports test statistics for significant differences between the participant groups and comparison groups. The first of these columns shows that Participant group I and Comparison group I are significantly different with respect to each of the variables gender, age, number of children, education and firm size. Replacing Comparison group I by Comparison group II removes the significant differences with regard to gender and age, but the differences for number of children, education and firm size remain significant. When we compare Participant group I with Comparison group III, there appear to be no significant differences with respect to age, education, firm size and number of children. Only for gender do we observe a significant difference at the 10% level. This suggests that a random event refraining someone from attending a training course is more likely to occur for women than for men. As examples of such random events the questionnaire refers to family circumstances and illness. While both events are arguably random, it is not surprising that they affect women slightly more than men: women are more often ill than men, and (at least in the Netherlands) there is still a tendency for women to bear a larger share of family responsibilities than men do.

The last three columns repeat the same exercise but now with Participant group I (all trained workers) replaced by Participant group II (workers who attended exactly one training/course). The results are very similar to those in the previous three columns. The most important difference is that now Participant group II and Comparison group III are no longer different with respect to their gender composition. This indicates that men and women have the same probability that a random event allocates them to Comparison group III rather than to Participant group II. In part the results on the tests are driven by sample size, but it is important to note that participants and comparisons do actually become increasingly similar when moving away from Comparison group I. This is especially the case for the important dimension on which they differ most, namely education.

The questionnaire also asked the respondents who attended a training course about the characteristics of this course. For a number of these training characteristics these questions were also asked to the respondents who wanted to receive training but did not do so. For instance, respondents who attended a course were asked who provided the course, while respondents who wanted to receive training but did not do so were asked who would have provided training. Such questions were asked with respect to the type of training, the provider of training, who paid the direct costs of training and whether training (would have) resulted in a certificate. Table IV presents the descriptive statistics of these characteristics for Participant group I, Comparison group II and Comparison group III. For respondents in Participant group I who followed more than one course, the answers relate to the first course they mention.

There are significant differences between the characteristics of the first training attended by respondents in Participant group I and the characteristics of the training which respondents in Comparison group II wanted to attend. These differences are in terms of the training type, provider and the party that pays the direct costs. Fewer differences are present when Participant group I and Comparison group III are compared. We still see significant differences

Table IV. Characteristics of training per participant and comparison group

	Participant	Comparison	
	I (1)	II (2)	III (3)
Type			
Foreign language	0.02	0.07	0.07
Safety	0.12	0.06	0.09
IT	0.17	0.18	0.18
Management	0.12	0.12	0.14
Communication	0.05	0.02	0.01
Marketing	0.03	0.02	0.01
Finance and administrative	0.05	0.05	0.02
Other occupation-related	0.28	0.26	0.29
Other	0.15	0.21	0.18
Provider			
Commercial organization	0.35	0.41	0.33
Employer	0.16	0.20	0.29
Sector/branch	0.08	0.10	0.11
Higher education Institute	0.09	0.04	0.01
Vocational school	0.05	0.01	0.02
Supplier	0.05	0.12	0.12
Other	0.22	0.12	0.11
Finance			
Employer	0.78	0.68	0.86
Employee	0.15	0.25	0.12
Both	0.05	0.07	0.02
Other	0.03	—	—
Certificate	0.76	0.80	0.83
χ^2 tests of equality (Pr > |t|)		PI vs. CII	PI vs. CIII
Type		31.8 (0.001)	14.0 (0.098)
Provider		40.5 (<0.001)	25.2 (0.002)
Finance		24.6 (<0.001)	4.8 (0.226)
Certificate		1.7 (0.196)	1.4 (0.229)

on training type and provider although all chi-square statistics drop substantially. Comparisons and participants, however, become remarkably similar on finance. This is important given that the focus of our analysis is on wage returns, which in turn depends on sharing of the training costs.

The results in this section suggest that, with respect to respondents' observable characteristics, Comparison groups II and III are more comparable to the two participant groups than Comparison group I. Moreover, the courses actually received by Participant group I and the courses which respondents in Comparison group III wanted to attend are more comparable than those of Comparison group II. This does not prove that Comparison group III is identical to a real randomly generated comparison group, but it is an indication that Comparison group III is more appropriate than Comparison groups I and II.

4. ESTIMATION RESULTS

Table V shows the coefficients of training in log wage equations for different combinations of participant and comparison groups, with and without control variables. The full set of variables in Table II are used as control variables (including age squared). Results are presented using mean regression (OLS) as well as using median regression (quantile regression). Mean regression is the standard method employed in the training literature. We also present results based on median regression to investigate the sensitivity of our results for the presence of outliers. We first discuss the results based on mean regression.

Without controls we find a log wage difference of 0.172 between Participant group I and Comparison group I. Adding controls reduces this difference to 0.113. Repeating this for Participant group II instead of Participant group I produces somewhat lower point estimates. But in all cases the wage differential between trained and untrained workers remains very substantial and is highly significant. This is in accordance with estimates of the effects of training incidence reported in other studies (see Section 1).

Table V. Effect of training on wages for different combinations of participant and comparison groups and control variables

	Type of regression	Control variables	Comparison group I			Comparison group II			Comparison group III		
			Coef. (1)	SE (2)	*p*-Value (3)	Coef. (4)	SE (5)	*p*-Value (6)	Coef. (7)	SE (8)	*p*-Value (9)
Participant I	Mean	No	0.172	(0.030)	[<0.001]	0.134	(0.042)	[0.001]	0.053	(0.049)	[0.281]
		Yes	0.106	(0.025)	[<0.001]	0.081	(0.034)	[0.018]	0.009	(0.055)	[0.869]
	Median	No	0.149	(0.028)	[<0.001]	0.097	(0.051)	[0.057]	0.051	(0.069)	[0.457]
		Yes	0.073	(0.016)	[<0.001]	0.059	(0.023)	[0.010]	0.028	(0.038)	[0.463]
N			2166			1270			1098		
Participant II	Mean	No	0.125	(0.035)	[<0.001]	0.087	(0.046)	[0.058]	0.006	(0.052)	[0.913]
		Yes	0.095	(0.029)	[0.001]	0.063	(0.036)	[0.086]	−0.010	(0.054)	[0.857]
	Median	No	0.116	(0.030)	[<0.001]	0.064	(0.040)	[0.107]	0.018	(0.048)	[0.705]
		Yes	0.076	(0.022)	[0.001]	0.063	(0.027)	[0.020]	0.033	(0.048)	[0.496]
N			1727			831			659		

Note: Controls are age, age squared, number of children, dummies for gender, level of education, firm size, tenure, being single, migrant and temporary job. Standard errors in parentheses; *p*-values in square brackets. Estimations use sample weights.

When we replace Comparison group I by Comparison group II the point estimates become somewhat smaller, but for both participant groups, with and without control variables, the training premium remains very substantial. As the number of observations in the comparison group reduces greatly (from 1145 to 249), the estimate is less precise but in all cases the coefficient differs significantly from zero.

This picture changes dramatically when Comparison group III serves as the comparison group; see column (7). In all cases the point estimate is reduced by a factor of three or more, and in none of the cases do we find a training premium significantly different from zero. This is particularly evident when considering the results for Participant group II, which is arguably the most appropriate treatment group given the construction of Comparison group III.

One might argue that our research design does not permit identification of these effects because of the limited sample size in Comparison group III. It should be noted, however, that increasing sample sizes to conventional numbers (as, for example, to the size of Comparison group I) would still not give us enough precision to identify effects of this size since the standard errors would only go down from 0.05 to 0.03. If the intervention is modest its effects are difficult to identify.[6]

To be more precise about the difference between the treatment effect based on Comparison groups I and III, we re-estimated the equations with the full set of control variables from column (1) and added the dummy for belonging to Comparison group III.[7] For Participant group I this shows a difference in log wage rates of 0.101 (SE 0.059) between individuals in Comparison group III and those who are in Comparison group I but not in Comparison group III. For Participant group II this difference is 0.103 (SE 0.057). These differences are statistically significant at the 10% level.

A possible concern with the results discussed above is that they are driven by a few outliers. For instance, if some of the 77 observations in Comparison group III have very high wages, this may lead to an underestimation of the true effect of training in column (7). To examine the sensitivity of the results, Table V also presents estimates based on median regression. The median regression results are fairly close to the mean regression results, indicating that the findings are not biased due to a few outliers.[8]

The approach in this paper builds on the availability of survey questions about reasons for non-participation that are normally not available. Without this information, researchers who want to purge the naive OLS estimates from selection bias have to rely on a parametric selection model or on some sort of matching. It is therefore interesting to apply these alternative methods to our data and compare the results with those presented in Table V.

When we apply a Heckman-type selectivity correction without imposing any exclusion restriction, hence identifying on functional form only (as in Lynch, 1992, and Veum, 1995), we get an estimate of 0.122 (SE 0.056). Using number of children and the dummy for single-status as

---

[6] Simulations show that when randomly drawing 77 observations from Comparison group I, in over 90% of the cases the estimated effect is larger than the point estimates reported in column (7) of Table V. Fewer than 11% of the simulations return a *p*-value less than that reported for Comparison group III.

[7] Note that this constrains the coefficients on the control variables to be the same in the regressions based on Comparison groups I and III. This has no consequences for our inference since the estimated difference is almost identical to the difference between participants and people in Comparison group I reported in Table V.

[8] Given the heterogeneity in private sector training, it is not clear how to interpret these results in terms of treatment effects. This would require quite strong assumptions on treatment assignment (that it is rank preserving). These additional assumptions are not necessary when estimating average effects as we do, since this involves linear operations.

variables included in the training equation but excluded from the wage equation, the estimate of the training return increases to 0.190 (SE 0.067). Using instead firm size as excluded variable the estimate increases even further to 0.344 (SE 0.116). When we perform nearest-neighbor matching where we match non-participant to participants on all the variables that are used in Table V as control variables, we obtain an estimate of 0.083 (SE 0.033). Matching on the propensity score gives an estimated return of 0.098 (SE 0.033). Note that both matching estimates are very close to the OLS estimate of 0.106 (SE 0.025) reported in Table V.

These alternative estimates can, of course, be criticized for their absence of a credible exclusion restriction or the breakdown of the conditional independence assumption in the case of matching, which are exactly the problems that motivated the approach in this paper.

5. CONCLUSION

Estimating returns to private sector training has turned out to be a very challenging research program. It is very difficult to come up with instruments, and no one has found a natural experiment. Consequently the literature has been relying on fixed-effects methods for the last two decades and estimated returns tend to be high, often several orders of magnitude higher than returns to schooling.

The main contribution of this paper lies in proposing an alternative approach to estimate the wage returns to private sector training that can be applied using a single cross-section. The idea is to restrict the group of untreated individuals to those who were willing to receive training but who did not do so due to some random event. Restricting the comparison group to those who were willing to participate eliminates biases due to self-selection of workers. Restricting the group of non-participating 'applicants' to those who did not participate due to some random event subsequently eliminates biases due to the selection process of firms.

The appropriateness of this newly created comparison group is corroborated by the similarity of this comparison group and the participant group in terms of a number of observed characteristics. Moreover, the courses that members of the comparison group wanted to attend and the characteristics of the courses actually attended by members of the participant group are also more comparable in terms observed training characteristics.

For our preferred specification, applying this approach leads to a reduction of the wage return to training from 10% to 17% (depending on covariates included and the exact participant group) to estimates close to zero for a training course with a median duration of 40 hours.

Recent theoretical contributions (Acemoglu and Pischke, 1999; Stevens, 1994) have emphasized the importance of imperfections in training markets. This literature shows that market imperfections give employers incentives to contribute more to the general training of their employees. The driving force behind this result is that market imperfections allow firms to capture part of the returns of training that is (at least partly) general. This result has implications for the literature that studies wage returns, namely that these will tend to be smaller (*ceteris paribus*) when market imperfections are more important. Market imperfections cannot only explain differences in training practices between countries but also (i) differences in the level of wage returns between countries, and (ii) wage returns that are small. Exploration of these links is an important area for future research that studies wage returns.

ACKNOWLEDGEMENTS

REFERENCES

Acemoglu D, Pischke J-S. 1999. The structure of wages and investment in general training. *Journal of Political Economy* **107**(3): 539–572.

Barron JM, Black DA, Loewenstein MA. 1993. Gender differences in training, capital, and wages. *Journal of Human Resources* **28**(2): 343–364.

Bartel AP. 1995. Training, wage growth, and job performance: evidence from a company database. *Journal of Labor Economics* **13**: 401–425.

Bell S, Orr L, Blomquist J, Cain GG. 1995. *Program Applicants as a Comparison Group in Evaluating Training Programs*. WE Upjohn Institute for Employment Research: Kalamazoo MI.

Blundell R, Dearden L, Meghir C. 1996. *The Determinants and Effects of Work Related Training in Britain*. Institute of Fiscal Studies: London.

Booth A. 1993. Private sector training and graduate earnings. *Review of Economics and Statistics* **75**(1): 164–170.

Fougère D, Goux D, Maurin E. 2001. Formation continue et carrières salariales: une évaluation sur données individuelles. *Annales d'Economie et de Statistique* **62**: 49–69.

Frazis H, Loewenstein MA. 2005. Reexamining the returns to training: functional form, magnitude, and interpretation. *Journal of Human Resources* **15**(2): 453–476.

Goux D, Maurin E. 2000. Returns to firm-provided training: evidence from French worker–firm matched data. *Labour Economics* **7**: 1–19.

Greenhalgh C, Stewart M. 1987. The effects and determinants of training. *Oxford Bulletin of Economics and Statistics* **49**: 171–189.

Heckman J, LaLonde R, Smith J. 1999a. The economics and econometrics of active labor market programs. In *Handbook of Labor Economics*, Vol. 3A, Ashenfelter O, Card D (eds) Elsevier Science: Amsterdam; Ch. 31.

Heckman JJ, LaLonde R, Smith J. 1999b. The economics and econometrics of active labor market programs. In *Handbook of Labor Economics*, Vol. 3A, Ashenfelter OC, Card D (eds). Elsevier Science: Amsterdam; Ch. 31.

Lengermann P. 1999. How long do the benefits of training last? Evidence of long term effects across current and previous employers. *Research in Labor Economics* **18**: 439–461.

Loewenstein M, Spletzer J. 1999a. General and specific training: evidence and implications. *Journal of Human Resources* **34**(4): 710–733.

Loewenstein MA, Spletzer JR. 1999b. *Formal and Informal Training: Evidence from the NLSY*, Vol. 18. JAI Press: Greenwich, CT; 403–438.

Lynch L. 1992. Private sector training and the earnings of young workers. *American Economic Review* **82**(1): 299–312.

Parent D. 1999. Wages and mobility: he impact of employer-provided training. *Journal of Labor Economics* **17**(2): 298–317.

Pischke J-S. 2001. Continuous training in Germany. *Journal of Population Economics* **14**: 523–548.

Stevens M. 1994. A theoretical model of on-the-job training with imperfect competition. *Oxford Economic Papers* **46**: 537–562.

Veum JR. 1995. Sources of training and their impact on wages. *Industrial and Labor Relations Review* **48**(4): 812–826.