EVALUATING THE EFFECT OF TAX DEDUCTIONS ON TRAINING

Edwin Leuven
Department of Economics - Universiteit van Amsterdam
Hessel Oosterbeek
Department of Economics - Universiteit van Amsterdam, and Tinbergen Institute

Abstract. Due to a tax law implemented in 1998, Dutch employers can claim an extra tax deduction when they train employees aged 40 years or older. This causes a discontinuity in a firm's cost of training an employee. This discontinuity is exploited to identify two effects: the effect of the tax deduction on training participation, and the effect of training participation on wages. The estimations show that the training rate of workers just above 40 is about 15-20 percent higher than the training rate of workers just below 40. This difference cannot be attributed to the stimulating effect of the tax deduction on the participation rate of older workers. Moreover, the short-term spill-over effects on workers younger than 40 are so substantial that the net effect of the age-dependent tax deduction is negative.

1. Introduction

Underinvestment in work-related training may occur for several reasons. With regard to general training, the so-called poaching externality is often mentioned as a source of underinvestment (see for example Stevens 1996). Regarding specific training, underinvestment is likely to occur because of the hold-up problem (Williamson 1985). As it is widely believed that workers' skills levels are important for the economic development of a country, governments typically regard such underinvestment as a problem and design policies to combat it. Training levies have been introduced in several countries as one of such policies. Another way to stimulate investment in work-related training is to create facilities for the deduction of training costs from taxable profits.

In 1998 the Dutch government implemented a new tax law which introduces three tax deductions for firms' expenditures on work-related training. The first deduction is general and gives all firms an extra deduction of their training expenditures from their taxable profits. (Training expenditures are already deducted when profits are calculated.) The second deduction, gives all firms spending less than a specified amount on worker training

---

Date: This version October 2000 (first draft June 2000).
We thank the following people for their comments: Joshua Angrist, Bas van der Klaauw, Mirjam van Praag, and Jim Spletzer. The usual disclaimer applies. The authors gratefully acknowledge financial support from NWO priority program 'Scholar' and the Max Goote Kenniscentrum for supplying us with the data.

an additional deduction of their training expenditures. The third applies to training costs made for employees aged 40 years or older. Firms are allowed to deduct an additional 40 percent of these costs.

The first deduction is meant to stimulate training participation in general, the second intends to stimulate worker training in small and medium sized firms, while the purpose of the third deduction is to enhance training participation of older workers. The last two deductions reflect the many empirical results which show that training rates increase with firm size and decrease with a worker's age (see for example: Barron et al. 1987, Lillard and Tan (1992), Royalty (1996), and Leuven and Oosterbeek (1999)).

The structure of the age-dependent tax deduction is discontinuous at age 40. All workers younger than 40 are excluded from this additional deduction, while all workers aged 40 or older are included. This structure constitutes a perfect example of a so-called "regression discontinuity (RD) data design" (cf. Campbell 1969). In this paper this feature is used to identify two effects: (i) the effect of the age-dependent tax deduction on the probability that a worker participates in training, and (ii) the effect of training participation on wages.

A maintained assumption in most (quasi-)experimental studies is the absence of external effects of the treatment. That is: there are no spill-overs from the treatment to the controls (or to the treated). As Philipson (2000) points out, in many cases this assumption is unlikely to be met and ignoring this can seriously bias the evaluation results. We explicitly take this into account in our estimations and show that this substantially alters the conclusions. External effects, in the sense that the control group is affected by the treatment, can take on two forms in the present application. Firstly, the treatment might lead to postponement of training participation until the worker is 40. Secondly, the employer might substitute training participation of a younger worker by training participation of an older worker. We propose a method to identify the combined magnitude of these external effects. This method requires data from the period before the introduction of the new tax law, and the identifying assumption that young workers (below 30) are not affected by the externality.

The standard RD estimate shows that training participation is substantially higher for workers just above 40 years than for workers just below 40 years. As it turns out, however, the identification of external effects is crucial for the interpretation of this finding. Using data from two cross-sections to account for external effects, it is found that the estimated effect is almost entirely due to a training dip in the control group.

Finally, the discontinuity of training participation at the age of 40 is used to create an instrumental variable which enables us to (locally) identify the causal effect on training participation on wages. As it turns out, the effect is not significantly different from zero. This finding does not depend on whether the estimation procedure takes the endogeneity of training into account.

One would also be interested in the effects of the other two tax deductions on training incidence. Unfortunately there are some serious problems involved with their evaluation. Regarding the general tax deduction, the problem is that this deduction applies to all workers and hence there is no cross-sectional control group. An alternative would be to compare training incidence over time, but this requires comparable time series data on training participation. Such data have not been collected for the Netherlands. And even if that would have been the case, data for a fairly long period would be needed as there are numerous other factors that potentially affect training incidence. The problem with evaluating the measure targeted to smaller firms is that it requires information from the tax files of individual firms. Since such information is confidential in the Netherlands, it would be very hard (or even impossible) to obtain it.

The analysis proceeds as follows. The next section provides more details about the new tax law. Section 3 discusses the strategy for the identification of causal effects. Section 4 considers the spill-over effects. Section 5 and 6 present the data and discuss the empirical findings. Finally, section 7 concludes.

2. The new tax law

January 1, 1998 a new tax law was introduced which gives firms the opportunity of additional tax deductions of their training expenditures. Three parts are distinguished. The first part is general and gives all firms an extra deduction of 20 percent of their training expenditures. The second part applies to firms that spend less than 250,000 DLG. These firms can deduct 40 percent instead of 20 percent of the first 60,000 DLG they spend on training.[1] Finally, all expenditures relating to the training of workers aged 40 years or older permit an additional deduction of 40 percent. The maximum deduction for a firm is 5,000,000 DLG.

Not all training costs are deductible. First of all it must concern the training of an employee and/or owner of the firm. In addition, the training must be relevant for the current function of the trainee. If the firm has its own training department, only the time employees (trainers) spend teaching is deductible. This means that preparation, development, and other overhead costs are not deductible. Finally, it applies only to formal training and only direct training costs qualify for tax deduction; that is: opportunity costs in the form of productivity foregone are excluded.

A numerical example illustrates the working of the new tax law. Consider a firm that spent 200,000 DLG on training expenditures during a fiscal year. Half of these

---

[1] This may have perverse effects on firms' incentives to offer their workers training. Firms that would otherwise spend slightly more than the threshold amount on training now have an incentive to reduce their (reported) training expenditures. Since there is no information available regarding the distribution of firms' training expenditures it is hard to tell how serious this effect is.

expenditures concerns older workers. The firm's total tax deduction is now equal to: 200,000 (the total training costs) + 40,000 (20 percent of the total training costs) + 40,000 (40 percent of the costs of training older workers) + 12,000 (20 percent of the first 60,000 since the firm's total training expenditures do not exceed 250,000).

In the Netherlands profits are taxed at 35 percent. This implies that, in general, 42 percent of training expenses are subsidized (35 percent of 100 + 20 percent). If the trained worker is older than 40, 56 percent of the training expenses are subsidized (35 percent of 100 + 20 + 40 percent).

If $C^*(t)$ are before tax training costs then after tax training costs equal

$$C(t) = \begin{cases} (0.51 - 0.14 \cdot 1_{\{Age \geq 40\}}) \cdot C^*(t) & \text{if } C^*(t) < 60,000 \\ (0.58 - 0.14 \cdot 1_{\{Age \geq 40\}}) \cdot C^*(t) - 4200 & \text{if } 60,000 \leq C^*(t) < 250,000 \\ (0.58 - 0.14 \cdot 1_{\{Age \geq 40\}}) \cdot C^*(t) & \text{if } C^*(t) \geq 250,000 \end{cases}$$

Here $1_{\{X\}}$ is an indicator variable taking on the value 1 if X is true and zero otherwise. Hence, the age-dependent tax deduction leads to a reduction of the marginal training costs of older workers of 14 percent.

3. Regression-discontinuity (RD) design

As stated in the introduction, this paper provides estimates of the effect of the tax policy on the training participation rate of older workers and, in addition, estimates the causal effect of training on wages. In both cases the fact is used that the new tax law creates a discontinuity at the age of 40. In the first case, being 40 years or older completely determines whether a worker qualifies for the treatment of the extra tax deduction; this treatment in turn affects the probability of training. In the second case, being 40 years or older is expected to increase the probability that someone is exposed to the treatment of training; this treatment in turn may affect wages. These are both applications of the regression-discontinuity method. The first case is an example a "sharp" regression discontinuity (RD) data design, while the second case is an example of a "fuzzy" design (cf. Campbell 1969). The two designs are presented following Hahn et al. (2000), who discuss identification issues, but the notation used is specific to the application in this paper.

The RD design exploits a known discontinuity in the treatment assignment to identify the treatment effect. In the sharp design, assignment to treatment depends in a deterministic way on a variable a_i with a known discontinuity at point $\overline{a}$ (in our case a_i is the age of worker i). This is the case for the assignment of workers to the treatment "age-dependent tax deduction". All workers of 40 years and older are treated, while all younger workers are not treated. Using d_i as indicator for assignment to the deduction treatment, the assignment rule is

$$d_i = \begin{cases} 1 \text{ if } a_i \geq \overline{a} = 40 \\ 0 \text{ otherwise} \end{cases}$$

The outcome is the probability of receiving training (t_i) and can be described as follows

$$E[t_i] = \alpha + \beta d_i$$

where $\alpha \equiv E[t_{0i}]$ is the training probability without extra tax deduction, and $\beta \equiv E[t_{1i}] - E[t_{0i}]$ is the change in training probability due to the extra tax deduction, the (common) treatment effect.

If there is no reason to believe that persons close to $\overline{a}$ are different then comparing persons just below this threshold with persons just above will give an unbiased estimate of the treatment effect:

$$\beta = t^+ - t^- \tag{1}$$

where $t^+ \equiv \lim_{a \downarrow \overline{a}} E[t|a]$ and $t^- \equiv \lim_{a \uparrow \overline{a}} E[t|a]$. The major identifying assumption is that there are no other discontinuities around $\overline{a}$. Unlike in a structural model and/or standard instrumental variable analysis, a_i may be correlated with the outcome variable t_i in which case assignment will not be random and a simple comparison of treated and non-treated will give a biased estimate of the treatment effect.

If training is taken as the treatment and the employee's (log) wage as the outcome, the setup corresponds to a fuzzy design. In a fuzzy design, assignment to treatment is not deterministic but probabilistic because it may depend on unobserved factors as well. The probability of treatment therefore becomes a function of a_i with a discontinuity at $\overline{a}$

$$Pr(t_i) = f(a_i, 1\{a_i \geq \overline{a}\})$$

As above, the outcome can be written as follows:

$$E[w_i] = \omega + \gamma t_i$$

where $\omega \equiv E[w_{0i}]$ is the (log) wage rate without training, and $\gamma \equiv E[w_{1i}] - E[w_{0i}]$ is the change in (log) wages due to training. It can be shown (again under the assumption of a common treatment effect) that γ can be identified by

$$\gamma = \frac{w^+ - w^-}{t^+ - t^-} \tag{2}$$

Because of the discontinuity at $\overline{a}$ the denominator does not equal zero. This formula is a local version of the Wald estimator and shows that RD is an IV estimator (this was first noted by Van der Klaauw 1996, but see also Angrist and Lavy 1999b). Note, however, that RD, contrary to IV, does not require that w_i, the outcome of interest, is independent of a_i, therefore the traditional exclusion restriction does not apply. All that is needed is the local continuity assumption (which is in fact an exclusion restriction regarding the discontinuity), and well defined limits in (2).

4. External treatment effects

The framework presented above (implicitly) assumes that members of the control group are not affected by the treatment, and that members of the treatment group are only affected by their own treatment and not by the treatment of others. In others words: it is assumed that external effects play no role. In many applications, this assumption can be disputed. An example in place concerns the estimation of returns to schooling using changes in compulsory schooling laws as an instrumental variable (e.g. Harmon and Walker (1994)). Those who were born just before the year that would include them in the treatment group of, say, a rise in the compulsory schooling age, belong - as an external effects of the change - to a group of low skilled people with relatively few competitors. This supply condition may increase their wage level relative to what their wages would have been in the absence of the compulsory schooling law. Since this wage level is the wage level of the control group, this mechanism would lead to a downward bias of the estimated effect of an extra year of schooling.

Philipson (2000) discusses the example of evaluating the effects of HIV-prevention trials. For such trials, external effects emerge when the infection rate of the control group in a trial depends on the share of treated. Another example mentioned by Philipson are spill-overs from R&D subsidies. Consider the extreme case where all firms that do not receive the treatment of the subsidy are, as a result of knowledge spill-overs, just as innovative as the firms that belong to the treatment group. Naive program evaluation would then indicate that R&D subsidies have no effect on firms' innovations while in reality the effects exceed the purely private benefits.[2]

The age-dependent tax deduction studied in this paper is another example of a treatment that might have spill-over effects to the (quasi-)control group. The reduction of training costs of older workers makes it relatively more expensive to train workers younger than 40. This might induce two different reactions from firms. The first is substitution.

---

[2] Philipson (2000) proposes a two-stage randomization scheme to disentangle the private effects and the external treatment effects. In the first stage the units within which the external effects occur are sampled and randomly assigned to different intensities of treatment. In the second stage, individuals within these units are randomly assigned to the treatment and control groups. This assumes the ideal situation in which the researcher can develop the experimental design. Most often this is not the case, and other sources of identification are needed.

An employer who wants to train a worker in order to acquire some new skills or knowledge, and who, in the absence of the treatment, would send a young worker to training may now decide to send an older worker to training. This implies instantaneous substitution of training an older worker for training a younger worker. The training rate of younger workers would decrease while the training rate of older workers would simultaneously increase. The second reaction to the tax deduction may be the postponement of training. Knowing that (direct) training costs will fall once workers turn 40, firms may delay training until this date. This implies intertemporary substitution of training. In steady state this would also lead to a drop in the training rate of younger workers and an increase of the training of older workers.

Although the consequences of the two mechanisms seem fairly similar, there are two differences. The substitution mechanism leads to an immediate increase of the training rate of older workers. For the postponement mechanism it will take some time to lead to an increase of the training rate of older workers (they first have to turn 40) out of steady state. Furthermore, while the substitution mechanism may lead to an increase of the training rates of all workers older than 40, the postponement mechanism leads only to an increase of the training rate of workers who just turned 40.

4.1 Identification

In the sharp design the causal effect of the policy can be estimated through equation (1). This does not tell us, however, to what extent the estimate reflects a decrease of training for the younger group due to the negative spill-over or a net increase of training for the older group. The external effect of the intervention for younger workers ($\text{effect}_{<40}$) can be defined in the following manner:

$$\text{effect}_{<40} \equiv E_{99}[t|a < 40, \pi = 1] - E_{99}[t|a < 40, \pi = 0] \tag{3}$$

where π equals unity if the policy is in place and zero otherwise. Similarly, the total effect of the intervention for older workers ($\text{effect}_{>40}$) is defined in a similar way:

$$\text{effect}_{>40} \equiv E_{99}[t|a > 40, \pi = 1] - E_{99}[t|a > 40, \pi = 0] \tag{4}$$

To identify these effects estimates are needed of training rates in 1999 in the absence of the policy: $E_{99}[t|a < 40, \pi = 0]$ and $E_{99}[t|a > 40, \pi = 0]$. One possible approach is to estimate a structural model based on the 1999 data and then change the policy parameters to calculate these counterfactuals. The disadvantage of this strategy is that it sensitive to the assumptions of the model. We will follow a more reduced-form strategy to create these counterfactuals. We will assume that, in the absence of the policy, the training-age profile would have shifted up or downwards relative to a pre-intervention base year:

Assumption 1a. $E_{99}[t|a, \pi = 0] = E_{94}[t|a, \pi = 0] + c$

or

Assumption 1b. $E_{99}[t|a, \pi = 0] = c' \cdot E_{94}[t|a, \pi = 0]$

Assumption 1a implies a parallel shift of the training-age profile between 1994 (the base year) and 1999, assumption 1b on the other hand reflects an equiproportionate shift of the training-age profile from 1994 to 1999. The reason for choosing 1994 as a base year is that for this year data are available that contain comparable information about the relation between training participation and age. Moreover, in this year the policy intervention was not in place, nor were there any plans for it, thereby ruling out possible anticipation effects.

Combining (3) and assumption 1a, it now follows that:

$$\text{effect}_{<40} \equiv E_{99}[t|a < 40, \pi = 1] - E_{94}[t|a < 40, \pi = 0] - c$$

Similarly (4) now becomes:

$$\text{effect}_{>40} \equiv E_{99}[t|a > 40, \pi = 1] - E_{94}[t|a > 40, \pi = 0] - c$$

The 'only' thing needed to identify $\text{effect}_{<40}$ and $\text{effect}_{>40}$ is an estimate of the shift parameter c. To identify c we need a control group that will be unaffected by the policy intervention. The best candidates for this are young workers, and the estimations will use the assumption that workers aged 25-30 are not affected by the intervention. This assumes that there are no spill-over effects from the treatment to workers in this age group. Reasons for this might be, first, that the period of postponement would be very long, and second, that workers below 30 are unlikely to be close enough substitutes to 40-year-old workers to replace them in a training spell.[3] Assumptions 1a and 1b are therefore augmented with the following assumptions:

Assumption 2a. $\hat{c} = \bar{t}^{99}_{25-30} - \bar{t}^{94}_{25-30}$

Assumption 2b. $\hat{c}' = \bar{t}^{99}_{25-30} / \bar{t}^{94}_{25-30}$

Estimators for the effects on the workers above and below 40 can now be derived. To illustrate the procedure under the two alternative assumptions of a parallel or a equiproportionate shift, they are written out for workers younger than 40 (changing the subscript to > 40 gives the other estimators). Under counterfactual assumptions 1a and 2a, the

---

[3] We tested for the sensitivity of the results for the use other brackets for the young age group. It turned out that this does not make much of a difference; see table A6 in the appendix.

following estimate of the effect on workers younger than 40 results:[4]

$$\widehat{\text{effect}}_{<40} = \bar{t}^{99}_{<40} - \bar{t}^{94}_{<40} - \hat{c} \tag{5}$$

Under assumptions 1b and 2b the following estimator results:

$$\widehat{\text{effect}}'_{<40} = \bar{t}^{99}_{<40} - \hat{c}' \cdot \bar{t}^{94}_{<40} \tag{6}$$

Finally, note that the availability of the 1994 data also provides us with an alternative, difference-in-differences, estimator for β as well:

$$\begin{aligned} \beta' &= (E_{99}[t|a > 40, \pi = 1] - E_{94}[t|a > 40, \pi = 0]) \\ &\quad -(E_{99}[t|a < 40, \pi = 1] - E_{94}[t|a < 40, \pi = 0]) \end{aligned}$$

comparing β to β' can serve as a test of the assumption that there are no other discontinuities at 40 that may affect training participation.

5. Data

This paper uses cross-section data from 1994 and 1999. The analysis rests mainly on the 1999 data which were collected from October to December 1999. 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, wages, and they also responded to a set of questions addressing the training activities they undertook in the 12 months prior to the interview. For the empirical analysis in this paper training participation is measured as an affirmative answer to the question whether the respondent participated in any training or course related to work or career during the last 12 months. Notice that for our purposes it is better to have data from a survey among workers than from administrative sources. While employers have reason to misreport the age of the workers who participated in training, workers have no such reason.

The analysis here focuses on employed persons aged 25 to 55. The age bracket is narrowed to avoid complications concerning education and retirement decisions. Restricting the sample this way ensures that most have made the transition from school to work, and in the Netherlands early retirement becomes an issue for workers over 55. The data contain information about the worker's age at the moment of the interview. This implies that workers aged 40 pose a problem for the analysis since it is not known whether or not there was a training opportunity they missed when they were 39. In the analyses they are therefore excluded.

---

[4] Appendix A1 gives the expressions for the variances of the two alternative effect measures.

To apply an RD design to evaluate the age-dependent tax deduction, one needs to compare observations just below the threshold age of 40 with observations just above that level. To secure that there are sufficient observations close to the point of discontinuity, there has been oversampling in the age brackets of 35-45 and 38-42. These are 2 respectively 5 years at each side of the point of discontinuity. The observations in the 35-45 bracket are used to create 5 so-called "discontinuity samples". Discontinuity sample±1 (abbreviated as DS±1) consists of the treatment group of 41-year-olds and the control group of 39-year-olds; DS±2 includes the treatment group of 41- and 42-year-olds and the control group of 38- and 39-year-olds. DS±3, DS±4 and DS±5 are defined in a similar manner.

Tables A1 and A2 in the appendix to this paper report some descriptive information of the sample. Table A1 gives the (weighted) sample means and standard deviations for the whole sample and for men and women separately. Table A2 provides some information about the differences between the treatment group and the control group in the various discontinuity samples. This shows that differences in observables are not important for narrow bands around 40 but that it is necessary to control for them as the window increases.

The 1994 data are needed to obtain estimates for c and c', defined in assumptions 1a-b. These data were collected as part of the Dutch wave of the so-called International Adult Literacy Survey (see Leuven and Oosterbeek (1999) for details). The important advantage of this dataset is that the phrasing of the training question in this survey is identical to the phrasing used in 1999. Barron et al. 1997 argue that existing measures of training show considerable differences and that this might be due to, among others, the survey instruments.

6. Evaluation of the effects of the tax deduction

6.1 Incidence

The probability of training participation conditional on age ($E[t_i|a_i]$) can be estimated using local linear regression (Cleveland 1979). Following Hahn et al. (2000), this is done by stratifying the sample on whether individuals were younger or older than 40 before doing the local linear regression.[5] Figure 1 plots, separately for 1994 and 1999, age against the estimated training participation rate. The squares in the figure show the actual average training rates by age. For 1999, the figure reveals a jump in these rates at age 40. The graph for 1994 shows that no such jump exists before the introduction of the

---

[5] Like Angrist and Lavy (1999a) the analysis uses the implementation in Stata of Cleveland's (1979) LOWESS estimator with Stata's default bandwidth of 0.8.

new tax policy. There appears to be a difference in employer behavior comparing 1994 to1999.

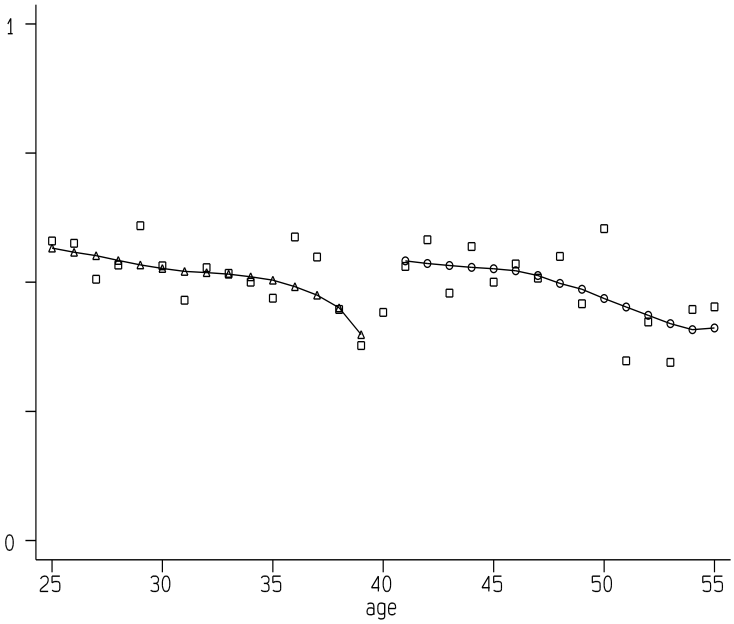
Table 1 provides estimates of the discontinuity for the 5 discontinuity subsamples around age 40. Column 1 gives the number of observations that fall in the specific DS-subsample. Although the sampling scheme was designed to oversample the age cohorts close to the point of discontinuity, DS±1 and DS±2 stratified by gender contain relatively few observations which may lead to a low level of precision of the estimates based on these subsamples. Columns 2 and 3 provide actual incidence rates for those younger and older than 40, respectively. Columns 4 to 7 report estimates of the jump shown in Figure 1 with differing combinations of controls for age, age squared, education, firm size and tenure. These estimates are obtained from OLS-regressions with a dummy variable for training participation as the dependent variable and a dummy for age older than 40 as (one of the) explanatory variable(s).

The estimates in column 4 give the difference in training incidence between different age groups without any controls. These estimates are therefore equal to the difference between the incidence rates in columns 3 and 2. For men and women together these estimates are all positive and decline when the range around age 40 is widened. The estimate that is conceptually closest to the one obtained using local linear regression is the estimate for DS±1, which is simply the difference in training rates between 41 and 39-year-olds. For the whole sample this gives us a statistically significant estimate of 17 percent. For males this estimate is about 19 percent and is on the border of significance. For women the estimated effect of the training deduction equals 13 percent, but lacks precision. These results are evidence of the working of the tax intervention since in the absence of the intervention, incidence rates would be decreasing in age and the estimates therefore negative.[6]

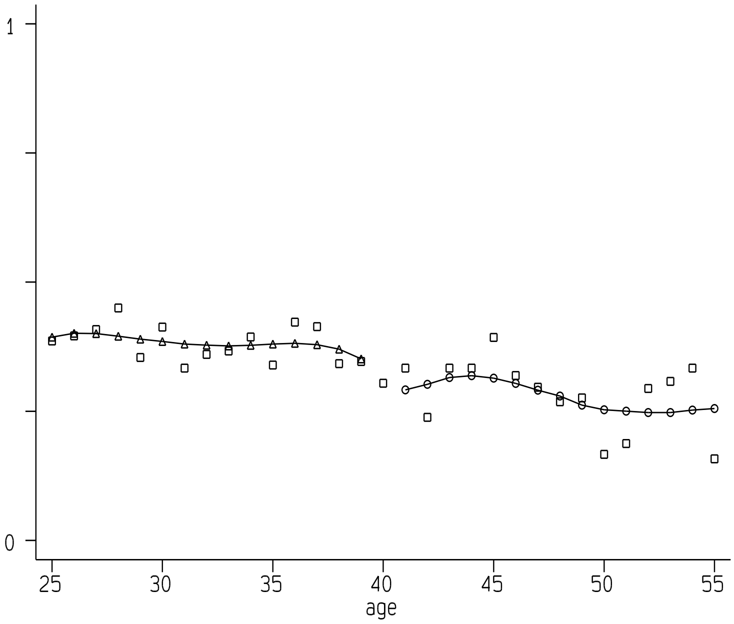
Column 5 adds controls for education, firm size and tenure to the regression equations. The estimates remain virtually identical to those in column 4. Columns 6 and 7 are based on regressions that are analogous to those in columns 4 and 5, but now age and age squared are added as regressors. For the DS±1 subsample this specification cannot be estimated as it is impossible to identify 3 age-related coefficients with only 2 age levels. For all other DS subsamples the estimates of the effect of the tax deduction, go up by a factor 2 or more relative to the estimates without controls for age and age squared. This reflects what is already apparent from Figure 1, namely that - except for the intervention - training incidence declines with age. As a result, the effect of the tax deduction is in the vicinity of 25 percent. For the DS±2 subsample the estimate is significant at the 10 percent level only, but notice that for this subsample only 4 age levels are used to identify

---

[6] Table A3 in the appendix reports comparable estimates for the 1994 sample, and shows that this is indeed the case.



(a) 1999



(b) 1994

Figure 1. Training incidence in 1999 and 1994

the coefficients of 3 age-related variables. For the separate subsamples of men and women, the picture is more erratic, but in almost all cases the point estimates are positive and point to substantial differences between the treatment groups and the control groups.

The results in table 1 reveal that if the width of the discontinuity sample increases it becomes important to control for differences in age. For DS±1 and DS±2 the estimated effects are significant without controls for age and age-squared (columns 4 and 5), while for DS±3 to ±5 the estimated effects are only significant when controls for age and age-squared are included (columns 6 and 7). The controls for age correct for the downward sloping training-age profile.

As an alternative the 1994-sample can be used to calculate difference-in-differences estimates. Table 2 presents the results. The first two columns give the difference in training rates between 1999 and 1994 for the control and treatment groups of the various discontinuity subsamples. For instance, the first row in the first column shows that between 1994 and 1999 there is an insignificant 4.7 percent fall in the training rate of 39-year-old workers (the control group in the DS±1 subsample). The final column gives the difference-in-differences estimates, that is: the difference in training rates between 1999 and 1994 for the treatment group minus the difference in training rates between 1999 and 1994 for the control group. The estimates in column 3 of table 2 are somewhat smaller than the estimates in the top panel of table 1, but clearly the effect is still very substantial, ranging from 10 percent for DS±5 to 20 percent for DS±2. Due to the small number of observations in 1994 the estimate for the DS±1 sample lacks precision, but the point estimate is in the same range.

6.2 Estimation of external effects

The results presented so far establish that there is a substantial difference in training rates between workers below age 40 and workers of 40 years or older. This difference can be attributed to the reduction of training costs of older workers caused by the age-dependent tax deduction. These results do not discriminate, however, between alternative interpretations of this difference. At one extreme, the whole difference can be interpreted as a stimulating effect of the tax deduction. At the other extreme, the difference can hide a drop in training rates of younger workers and a rise in training rates of older workers, both resulting from substitution and/or postponement. In the first case the net effect - in terms of an increase in training participation - equals the estimated difference. In the second case the net effect equals zero (or can be even negative in the short run).

Column 1 of table 3 provides estimates of the external effect for workers younger than 40 based on equation (5) and column 2 gives the equivalent estimate for workers older than 40. For all 5 discontinuity subsamples a negative effect for workers younger than 40

Table 1. Training participation rates discontinuity samples (DS), 1999

	N	< 40	> 40	Δ=(3)-(2)			
DS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Men & women							
±1	204	0.377	0.547	0.170	0.163		
				(0.069)**	(0.072)**		
±2	385	0.425	0.555	0.130	0.120	0.241	0.273
				(0.051)**	(0.051)**	(0.157)	(0.155)
±3	538	0.462	0.536	0.074	0.066	0.295	0.308
				(0.043)*	(0.043)	(0.109)***	(0.108)***
±4	683	0.492	0.543	0.051	0.039	0.237	0.254
				(0.038)	(0.039)	(0.088)***	(0.088)***
±5	840	0.486	0.535	0.048	0.013	0.164	0.177
				(0.035)	(0.035)	(0.077)**	(0.075)**
Men							
±1	107	0.403	0.590	0.188	0.190		
				(0.096)*	(0.101)*		
±2	185	0.493	0.570	0.077	0.087	0.434	0.518
				(0.074)	(0.077)	(0.224)*	(0.227)**
±3	260	0.516	0.577	0.061	0.071	0.241	0.287
				(0.062)	(0.064)	(0.153)	(0.154)*
±4	336	0.554	0.565	0.011	0.013	0.248	0.269
				(0.054)	(0.057)	(0.123)**	(0.125)**
±5	423	0.532	0.553	0.021	−0.005	0.127	0.120
				(0.049)	(0.051)	(0.108)	(0.108)
Women							
±1	97	0.341	0.470	0.130	0.138		
				(0.100)	(0.103)		
±2	200	0.342	0.533	0.191	0.206	−0.009	−0.012
				(0.070)***	(0.069)***	(0.220)	(0.213)
±3	278	0.397	0.472	0.075	0.067	0.383	0.442
				(0.060)	(0.059)	(0.155)**	(0.152)***
±4	347	0.418	0.506	0.088	0.074	0.205	0.266
				(0.054)	(0.053)	(0.127)	(0.124)**
±5	417	0.430	0.503	0.073	0.044	0.193	0.228
				(0.049)	(0.049)	(0.109)*	(0.105)**
Controls							
Age				*No*	*No*	*Yes*	*Yes*
Other				*No*	*Yes*	*No*	*Yes*

Notes: Controls denoted by “Other” are 7 education dummies, 5 firm size dummies, and tenure. Controls denoted by “Age” are age and age squared. Coefficients shown in columns 4 to 7 are the coefficients of a dummy in an OLS regression. Individuals aged 40 are excluded from the estimations. Standard errors in parentheses. * significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

Table 2. Difference-in-differences estimates

	$\Delta_{<40}$ (1)	$\Delta_{>40}$ (2)	Δ=(2)-(1) (3)
±1	−0.047	0.154	0.201
	(0.084)	(0.100)	(0.130)
±2	0.045	0.246	0.201
	(0.061)	(0.068)***	(0.091)**
±3	0.077	0.216	0.139
	(0.050)	(0.054)***	(0.073)*
±4	0.098	0.209	0.111
	(0.042)**	(0.047)***	(0.063)*
±5	0.100	0.194	0.095
	(0.038)***	(0.043)***	(0.057)*

Notes: Individuals aged 40 are excluded from the estimations. Standard errors in parentheses; * significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

is found. For DS±1 to DS±3 these negative effects are significantly different from zero, and range from 11 to 23 percent. This is a clear indication of the presence of negative spill-over effects. Without any spill-overs, workers younger than 40 would not be affected and the estimated effect would equal zero. The fact that for DS±4 and DS±5, the negative effect is not significantly different from zero further indicates that the negative spill-overs are mostly concentrated among workers with ages just below 40.

For workers older than 40 none of the estimated effects in column 2 is significantly different from zero (for the smallest discontinuity subsample the point estimate is even negative). This is a puzzling finding: if we find the negative spill-overs for the younger age group we also expect to find its mirror image in the form of a positive effect for the older age group. To test whether the intervention has indeed adverse effects, column 3 reports the sum of the two effects along with its standard deviation. For DS±1 it is found that this sum is significantly negative. This suggest that that the negative effect for younger workers is not compensated by a positive effect for older workers. From the perspective of the intention behind the age-dependent tax deduction – stimulating training participation of older workers – these findings are discouraging.

One explanation for this result possibly lies in the fact that the policy has only been implemented recently (January 1998), and that it may take firms some time to adjust their

Table 3. Estimates of external and direct effects

	$E_{99}[t \mid a] = E_{94}[t \mid a] + c$			$E_{99}[t \mid a] = c' \cdot E_{94}[t \mid a]$		
	$\text{effect}_{<40}$ (1)	$\text{effect}_{>40}$ (2)	(1)+(2) (3)	$\text{effect}_{<40}$ (4)	$\text{effect}_{>40}$ (5)	(4)+(5) (6)
±1	−0.232	−0.032	−0.264	−0.256	−0.040	−0.306
	(0.092)**	(0.104)	(0.139)*	(0.118)**	(0.134)	(0.179)*
±2	−0.140	0.061	−0.079	−0.142	0.093	−0.049
	(0.071)*	(0.075)	(0.103)	(0.092)	(0.090)	(0.129)
±3	−0.108	0.031	−0.077	−0.112	0.058	−0.054
	(0.062)*	(0.064)	(0.089)	(0.079)	(0.077)	(0.110)
±4	−0.087	0.024	−0.063	−0.096	0.045	−0.051
	(0.056)	(0.059)	(0.081)	(0.073)	(0.072)	(0.103)
±5	−0.086	0.009	−0.077	−0.091	0.026	−0.065
	(0.053)	(0.056)	(0.077)	(0.069)	(0.069)	(0.098)

Notes: t^{99}_{25-30}=0.567 (0.023), and t^{94}_{25-30}=0.382 (0.028). This gives us estimates for c=0.185 (0.037), and c'=1.493 (0.123). Standard errors are in parentheses; * significant at 10 percent level; ** significant at 5 percent level; *** significant at 1 percent level.

training policy to the new rules. If the spill-over to the younger age group is primarily due to postponement of training participation then we may observe in 1999 the postponement of training among the controls but not yet the catching up of the postponed training by the treated. Notice that this does not depend on the exclusion of the 40-year-olds from the analysis. As panel b of figure 1 shows, the training rate of 40-year-olds lies between the training rates of 39- and 41-year-old workers.

Columns 4-6 give alternative estimates based on equation (6). These give qualitatively fairly similar results: for younger workers the point estimates are almost the same, for older workers the point estimates are (1.5 to 2 times) larger, and for both age groups the standard errors are somewhat larger.

6.3 Wage effects

Previous studies have investigated wage effects of employer-provided training; examples include Bishop (1994), Lynch (1992), Pischke (1996), Lillard and Tan (1992). The basic problem in estimating these effects is that selection into training is likely to be non-random and might depend on non-observables that correlate with wages. Comparing the wages of trained workers with wages of non-trained workers either directly or through regressing wages on training will then give a biased estimate. Two solutions have typically

Table 4. Wage returns to training OLS and 2SLS estimates

	OLS						
	coef. (1)	s.e. (2)	N (3)	R^2 (4)	F d40=0 (5)	Pr $>F$ (6)	R^2_{d40} (7)
±1	0.053	(0.096)	149	0.42			
±2	-0.014	(0.069)	275	0.27			
±3	0.001	(0.056)	391	0.25			
±4	0.006	(0.048)	503	0.25			
±5	0.002	(0.044)	626	0.25			
25-55	0.030	(0.026)	1627	0.26			
	2SLS						
±1	...	...	...	...	...	...	...
±2	-0.495	(0.605)	275	0.13	3.030	0.083	0.091
±3	-0.315	(0.411)	391	0.19	8.089	0.005	0.010
±4	0.107	(0.410)	503	0.25	8.347	0.004	0.000
±5	0.283	(0.484)	626	0.20	5.574	0.018	0.003
25-55	-0.063	(0.460)	1627	0.25	5.987	0.014	0.075

Notes: Controls include age and age squared, 7 education dummies, 5 firm size dummies, (job) tenure, # children, migrant, female, temporary, single. Individuals aged 40 are excluded from the estimations. Standard errors in parentheses.

been implemented. If the researcher has access to longitudinal data it is possible to eliminate time-invariant individual characteristics through a fixed effects regression. This will eliminate the bias arising from selection on these time-invariant unobservables. A second solution would be to find instrumental variables for training; variables that affect wages only through training. As discussed in section 3, the discontinuity around age 40 can be used to identify the wage returns to training exploiting the fuzzy design of the data. This basically boils down to doing local instrumental variables estimation using an indicator variable for the discontinuity as an instrument.

Table 4 reports OLS and 2SLS estimates for the various DS samples. For all 5 DS samples, OLS estimates of the wage returns to training are not significantly different from zero. This is also the case for the entire sample of 25 to 55-year-old workers. All wage regressions include the covariates sets Age and Other (cf. Table 1). The same conclusion emerges from the 2SLS estimates. The point estimates are much larger in absolute size, but so are the standard errors and as a result none of the estimated wage effects differs significantly from zero.

Although not typical, the finding of a negligible wage return to company training participation is not uncommon. Two factors seem to matter in this respect: (i) returns seem to be lower in Europe than in the US, and (ii) most US studies estimate returns for younger workers whereas this paper builds on a sample of older workers. Examples of studies with low or zero returns are Booth (1993) for the UK, and Pischke (1996) for Germany. Lynch (1992), and Veum (1995) report returns to company training incidence not statistically different from zero for US (NLSY) data. Loewenstein and Spletzer (1999) find positive returns using more recent waves of the NLSY data.

The findings indicate that the zero wage return is not the result of any unobserved heterogeneity. It should be noted, however, that the IV estimates presented in table 4 identify the wage returns for a fairly specific group of workers and based on fairly specific training spells. For the the DS±1 subsample for example, the wage returns are identified by (see equation 2) the difference in wages between 41- and 39-year-old workers divided by their difference in training rates.

7. Conclusion

The new tax law implemented in the Netherlands in January 1998 causes a discontinuity in firms' direct costs of training as a function of the age of a worker. These costs are 14 percent lower for workers who are 40 years or older than for workers who are younger than 40. This feature of the tax law is exploited to identify two effects. Firstly, the effect of a 14 percent cost reduction on the probability that a worker participates in training. Secondly, the effect of training participation on the worker's wage rate. Since both effects are identified only locally, the participation effect of a cost reduction pertains to 40-year-old workers. Similarly, the wage effect is measured as the wage effect of training participation by a 40-year-old worker.

With regard to the effect of a cost reduction it is found that the training rate of workers somewhat older than 40 is about 15-20 percent higher than the training rate of workers who are slightly younger than 40. This difference cannot entirely be attributed to the stimulating effect of the cost reduction. On the contrary: The estimates of the spill-over effects on workers younger than 40 indicate that these spill-overs are so substantial that the net effect of the age-dependent tax deduction is negative. It cannot be ruled out, however, that this result is partly due to the fact that the data have been collected shortly (less than 2 years) after the policy has been implemented. Perhaps the data fully reflect the negative spill-over to the younger age group but do not reflect yet the catching up of postponed training by workers in the older age group. But even if the results in this paper reflect only short term effects, they show that spill-over effects can be very important and that neglecting such effects can seriously bias the policy evaluation. Had the potential of spill-over effects in this study been ignored, we would have concluded that the age-dependent tax deduction has increased training rates of older workers in the region 15-20 percent.

Regarding the effect of training participation on wages, both the OLS and 2SLS estimates show that they are not significantly different from zero. This finding indicates that the zero wage return is, at least for the group of workers with ages around 40, not the result of unobserved heterogeneity. It should be noted that the wage effects identified are also short term effects. This estimate relates wage rates measured at the end of 1999 to training participation that occurred at some time during the 12 months prior to the date of the interview. Perhaps it takes more time for wages to respond to training participation.

From the perspective of the policy-makers who intended to stimulate training participation of older workers the results in this paper are discouraging. More in general, reservations can be made regarding the effectiveness of training subsidies. Holzer et al. (1993) note that a potential risk of these subsidies is that these are merely windfall gains for the recipients and that the only result is a substitution of public for private spending. This concern is also relevant in the context of the tax rule studied in the current paper.

References

Angrist, J. D. and V. Lavy (1999a). New evidence on classroom computers and pupil learning. Working Paper Series 7424, NBER.

Angrist, J. D. and V. Lavy (1999b). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics 114*(2), 533–575.

Barron, J. M., M. C. Berger, and D. A. Black (1997). How well do we measure training? *Journal of Labor Economics 15*(3), 507–528.

Barron, J. M., D. A. Black, and M. A. Loewenstein (1987). Employer size: The implications for search, training, capital investment, starting wages, and wage growth. *Journal of Labor Economics 5*(1), 76–89.

Bishop, J. (1994). The impact of previous training on productivity and wages. In L. Lynch (Ed.), *Training and the Private Sector: International Comparisons*, NBER comparative labor markets series. Chicago: University of Chicago Press.

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

Campbell, D. (1969). Reforms as experiments. *American Psychologist 24*, 409–429.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association 74*(368), 829–836.

Hahn, J., P. Todd, and W. Van der Klaauw (2000). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, Forthcoming.

Harmon, C. and I. Walker (1994). Estimates of the economic return to schooling for the United Kingdom. *American Economic Review 85*(5), 1278–86.

Holzer, H. J., R. N. Block, M. Cheatham, and J. H. Knott (1993). Are training subsidies for firms effective? The Michigan experience. *Industrial and Labor Relations Review 46*(4), 625–636.

Leuven, E. and H. Oosterbeek (1999). Demand and supply of work-related training; evidence from four countries. *Research in Labor Economics 18*, 303–330.

Lillard, L. and H. Tan (1992). Private sector training: Who gets it and what are its effects? *Research in Labor Economics 13*, 1–62.

Loewenstein, M. A. and J. R. Spletzer (1999). Formal and informal training: evidence from the NLSY. *Research in Labor Economics 18*, 403–438.

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

Mood, A., F. Graybill, and D. Boes (1974). *Introduction to the Theory of Statistics.* McGraw-Hill.

Philipson, T. (2000). External treatment effects and program implementation bias. Technical Working Paper Series T0250, NBER.

Pischke, J.-S. (1996). Continuous training in Germany. Working Paper Series 5829, NBER.

Royalty, A. B. (1996). The effects of job turnover on the training of men and women. *Industrial and Labor Relations Review 49*(3), 506–521.

Stevens, M. (1996). Transferable training and poaching externalities. In A. Booth and D. Snower (Eds.), *Acquiring skills: market failures, their symptoms and policy responses.* Cambridge University Press.

Van der Klaauw, W. (1996). A regression-discontinuity evaluation of the effect of financial aid offers on college enrollment. Manuscript, New York University, Department of Economics.

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

Williamson, O. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting.* New York: Free Press.

A. Appendices

A.1 Expressions for the variances of effect

It is straightforward to derive the variance of (5):

$$var(\widehat{\text{effect}}_{<40}) = var(\bar{t}^{99}_{<40}) + var(\bar{t}^{94}_{<40}) + var(\bar{t}^{99}_{25-30}) + var(\bar{t}^{94}_{25-30})$$

assuming that $\bar{t}^{99}_{25-30}$, $\bar{t}^{94}_{25-30}$, $\bar{t}^{99}_{<40}$, and , $\bar{t}^{94}_{<40}$ are independent. Similarly the variance of (6) follows after some manipulation:

$$var(\widehat{\text{effect}}'_{<40}) = var(\bar{t}^{99}_{<40}) \cdot (\bar{t}^{94}_{<40})^2 + var(\hat{c}') \cdot (\hat{c}')^2 \cdot var(\bar{t}^{94}_{<40}) + var(\bar{t}^{94}_{<40}) \cdot var(\hat{c}')$$

where

$$var(\hat{c}') \approx \left(\frac{\bar{t}^{99}_{25-30}}{\bar{t}^{94}_{25-30}}\right)^2 \left(\frac{var(\bar{t}^{99}_{25-30})}{(\bar{t}^{99}_{25-30})^2} + \frac{var(\bar{t}^{94}_{25-30})}{(\bar{t}^{94}_{25-30})^2}\right)$$

again, assuming independence of $\bar{t}^{99}_{25-30}$, $\bar{t}^{94}_{25-30}$, $\bar{t}^{99}_{<40}$, and , $\bar{t}^{94}_{<40}$ (see f.e. Mood, Graybill, and Boes (1974) p. 180-81).

Table A1. Sample descriptives (weighted), 1999

	All		Men		Women	
	mean	s.d.	mean	s.d.	mean	s.d.
Age	38.6	(8.6)	39.1	(8.7)	37.8	(8.4)
Tenure (months)	115.7	(107.0)	133.4	(115.7)	89.7	(86.5)
Tenure in function (months)	79.5	(86.7)	87.1	(93.6)	68.4	(74.0)
# Children	0.97	(1.13)	0.95	(1.15)	1.00	(1.10)
ln(hours wage)	3.375	(0.58)	3.482	(0.56)	3.206	(0.59)
ln(weekly wages)	6.822	(0.74)	7.084	(0.56)	6.408	(0.80)
ln(hours working week)	3.440	(0.42)	3.612	(0.22)	3.189	(0.51)
Single	0.170	(0.38)	0.189	(0.39)	0.143	(0.35)
Migrant	0.067	(0.25)	0.067	(0.25)	0.067	(0.25)
Temporary	0.076	(0.27)	0.056	(0.23)	0.105	(0.31)
Female	0.405	(0.49)				
Education:						
Lower	0.036	(0.19)	0.038	(0.19)	0.033	(0.18)
Lower secondary: vocational	0.113	(0.32)	0.128	(0.33)	0.091	(0.29)
Lower secondary	0.121	(0.33)	0.094	(0.29)	0.160	(0.37)
Upper secondary: vocational	0.278	(0.45)	0.264	(0.44)	0.299	(0.46)
Upper secondary	0.063	(0.24)	0.063	(0.24)	0.063	(0.24)
Tertiary: Non university	0.258	(0.44)	0.264	(0.44)	0.249	(0.43)
Tertiary: University	0.110	(0.31)	0.126	(0.33)	0.087	(0.28)
Missing	0.021	(0.14)	0.022	(0.15)	0.018	(0.13)
Firmsize:						
1-10	0.106	(0.31)	0.087	(0.28)	0.135	(0.34)
11-50	0.217	(0.41)	0.205	(0.40)	0.236	(0.42)
51-100	0.119	(0.32)	0.127	(0.33)	0.107	(0.31)
101-200	0.126	(0.33)	0.152	(0.36)	0.088	(0.28)
¿200	0.391	(0.49)	0.405	(0.49)	0.371	(0.48)
Don't know	0.041	(0.20)	0.025	(0.16)	0.064	(0.24)
N	2326		1178		1148	
N (weighted)	2326		1383		943	

Table A2. P-values for equality of treatment and control groups in various DS samples

	DS±1	DS±2	DS±3	DS±4	DS±5
Tenure (months)	0.307	0.073	0.008	0.000	0.000
Tenure in function (months)	0.983	0.259	0.023	0.000	0.000
# Children	0.634	0.027	0.006	0.000	0.000
ln(hourly wage)	0.296	0.069	0.005	0.004	0.001
ln(weekly wage)	0.274	0.079	0.015	0.019	0.005
ln(hours working week)	0.503	0.608	0.732	0.642	0.511
Single	0.919	0.404	0.234	0.657	0.136
Migrant	0.543	0.353	0.528	0.830	0.514
Temporary	0.715	0.282	0.096	0.106	0.056
Female	0.448	0.350	0.115	0.047	0.016
Education	0.986	0.260	0.103	0.140	0.114
Firmsize	0.995	0.440	0.193	0.087	0.330

Notes: All reported p-values are from t-tests. For this purpose education is measured in years and firm size in number of employees.

Table A3. Training participation rates discontinuity samples (DS), 1994

	N	<40	>40	Δ =(3)-(2)			
DS	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Men & women							
±1	88	0.424	0.393	−0.031	−0.034		
				(0.108)	(0.111)		
±2	168	0.38	0.310	−0.070	−0.038	0.011	0.007
				(0.074)	(0.076)	(0.234)	(0.234)
±3	274	0.384	0.320	−0.064	−0.053	−0.043	−0.064
				(0.058)	(0.058)	(0.157)	(0.153)
±4	387	0.394	0.334	−0.061	−0.042	−0.057	−0.044
				(0.049)	(0.048)	(0.128)	(0.125)
±5	468	0.387	0.341	−0.046	−0.031	−0.08	−0.083
				(0.045)	(0.044)	(0.112)	(0.111)
Men							
±1	52	0.526	0.310	−0.216	−0.246		
				(0.133)	(0.137)∗		
±2	90	0.417	0.341	−0.076	−0.086	−0.527	−0.523
				(0.098)	(0.099)	(0.300)∗	(0.297)∗
±3	148	0.449	0.333	−0.116	−0.112	−0.148	−0.206
				(0.078)	(0.079)	(0.205)	(0.202)
±4	219	0.441	0.346	−0.095	−0.086	−0.169	−0.193
				(0.067)	(0.065)	(0.166)	(0.163)
±5	272	0.426	0.364	−0.062	−0.056	−0.207	−0.226
				(0.060)	(0.059)	(0.146)	(0.144)
Women							
±1	36	0.184	0.563	0.379	0.345		
				(0.164)∗∗	(0.188)∗		
±2	78	0.292	0.267	−0.025	0.01	1.078	1.075
				(0.113)	(0.130)	(0.350)∗∗∗	(0.387)∗∗∗
±3	126	0.268	0.303	0.035	−0.003	0.187	0.254
				(0.085)	(0.089)	(0.249)	(0.248)
±4	168	0.323	0.317	−0.006	0.003	0.165	0.22
				(0.072)	(0.074)	(0.203)	(0.205)
±5	196	0.326	0.302	−0.024	−0.016	0.176	0.203
				(0.066)	(0.068)	(0.179)	(0.179)
Controls							
Age				*No*	*No*	*Yes*	*Yes*
Other				*No*	*Yes*	*No*	*Yes*

Notes: See table 1.

Table A4. Wage returns to training OLS and IV estimates, men

	coef	se	N	R^2	F d40=0	Prob>F	R^2_{d40}
OLS							
25-55	0.022	(0.035)	851	0.26			
±1	-0.089	(0.154)	84	0.45			
±2	-0.074	(0.093)	139	0.33			
±3	-0.092	(0.072)	199	0.31			
±4	-0.056	(0.063)	260	0.27			
±5	-0.004	(0.053)	334	0.26			
2SLS							
25-55	-0.013	(0.642)	851	0.26	2.206	0.138	0.138
±1	...	...	...	...	...	...	...
±2	-0.187	(0.339)	139	0.32	5.904	0.016	0.004
±3	-0.237	(0.375)	199	0.30	3.455	0.064	0.003
±4	0.164	(0.447)	260	0.24	4.278	0.039	0.015
±5	0.696	(0.716)	334		1.411	0.236	0.000

Notes: See table 4.

Table A5. Wage returns to training OLS and IV estimates, women

	coef	se	N	R^2	F d40=0	Prob>F	R^2_{d40}
OLS							
25-55	0.037	(0.041)	776	0.20			
±1	0.093	(0.126)	65	0.62			
±2	-0.029	(0.117)	136	0.28			
±3	0.003	(0.093)	192	0.25			
±4	-0.032	(0.079)	243	0.26			
±5	-0.056	(0.079)	292	0.26			
2SLS							
25-55	-0.092	(0.678)	776	0.19	3.099	0.079	0.008
±1	...	...	...	...	...	...	...
±2	3.004	(6.098)	136		0.030	0.863	0.184
±3	-0.396	(0.705)	192	0.17	6.859	0.009	0.003
±4	0.063	(0.796)	243	0.25	4.694	0.031	0.012
±5	-0.011	(1.044)	292	0.26	4.371	0.037	0.007

Notes: See table 4.

Table A6. Sensitivity of the estimates of c and c' for the choice of age brackets

	$\hat{c}$	s.e.	$\hat{c}'$	s.e.
25-25	0.193	(0.089)	1.533	(0.291)
25-26	0.225	(0.064)	1.632	(0.233)
25-27	0.191	(0.052)	1.528	(0.181)
25-28	0.164	(0.044)	1.419	(0.137)
25-29	0.191	(0.040)	1.505	(0.134)
25-30	0.185	(0.037)	1.493	(0.123)
25-31	0.177	(0.034)	1.475	(0.114)
25-32	0.185	(0.032)	1.503	(0.109)
25-33	0.180	(0.030)	1.488	(0.101)
25-34	0.173	(0.028)	1.463	(0.093)
25-35	0.168	(0.027)	1.453	(0.089)