# Evaluating the Effect of Tax Deductions on Training

Edwin Leuven, *University of Amsterdam and Tinbergen Institute*

Hessel Oosterbeek, *University of Amsterdam and Tinbergen Institute*

Dutch employers can claim an extra tax deduction when they train employees older than age 40. This discontinuity in a firm's training cost is exploited to identify the tax deduction's effects on training participation and of training participation on wages. The results show that the training rate of workers just above age 40 is 15%–20% higher than that of workers just below age 40. This difference mainly results from the postponement of training and is not a stimulating effect of the measure. The two-stage least squares estimate of the wage effect of training is not statistically different from zero.

## I. 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, e.g., Stevens 1994). Regarding specific training, underinvestment is likely to occur because of

We gratefully acknowledge valuable comments from Joshua Angrist, Bas van der Klaauw, Wilbert van der Klaauw, Mirjam van Praag, and Jim Spletzer. The usual disclaimer applies. We gratefully acknowledge financial support from Nederlandse Organisatie Voor Wetenschappelijk Onderzoek (NWO) priority program "Scholar" and the Max Goote Kenniscentrum for supplying us with the data. Contact the corresponding author, Edwin Leuven, at e.leuven@uva.nl.

[*Journal of Labor Economics,* 2004, vol. 22, no. 2]

0734-306X/2004/2202-0008$10.00

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 (e.g., Brazil, France, Singapore, and Turkey) 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 that 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 an additional deduction of their training expenditures. The third applies to training costs made for employees ages 40 years or older. Firms are allowed to deduct an additional 40% 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 showing that training rates increase with firm size and decrease with a worker's age (see, e.g., Barron, Black, and Loewenstein 1987; Lillard and Tan 1992; Royalty 1996; Leuven and Oosterbeek 1999).

The structure of the age-dependent tax deduction is discontinuous at the age of 40. All workers younger than age 40 are excluded from this additional deduction, while all workers ages 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 article, 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 effects from the treatment on the controls. 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 treatment effects can take on two forms in the present application. First, the treatment may lead to postponement of training participation until the worker is age 40. Second, the employer may substitute training participation of a younger worker with training participation of an older worker. We propose a method to identify the combined magnitude of these external treatment effects. This method requires data from the period before the introduction of the new tax law, and the identifying assumption is that young workers (below age 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 treatment effects is crucial for the interpretation of this finding. Using data from two cross-sections to account for these 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 age 40 is used to create an instrumental variable that enables us to identify locally 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. While we tried our very best, we did not succeed in getting access to these data.

Obviously, the other two tax deductions are likely to increase training participation. As workers of different ages may be affected differently, these deductions may change the age-training profile. We argue that this does not affect the estimate of the local effect of the age-dependent tax deduction, nor does it affect the instrumental variable estimate of the wage effect of training. It may, however, weaken our strategy to identify the external treatment effects. Section III.C discusses how this issue is addressed.

The analysis proceeds as follows: Section II provides more details about the new tax law. Section III discusses the strategy for the identification of causal effects. This section briefly discusses the regression discontinuity design and its identifying assumptions. In addition, it discusses the approach taken in this article to take external treatment effects into account. In Section IV, the data are described. Section V presents the estimates of the effect of the tax deduction on training participation and the subsequent estimates of the external treatment effects. The wage effects of training are presented in Section VI. Finally, Section VII concludes.

## II. The New Tax Law

During the 1990s there was a growing awareness in the Netherlands (as elsewhere) that the skill level of the workforce is an important determinant of economic prosperity. In 1996 and 1997, the Dutch government launched broad discussions about the concept of lifelong learning, after which it came up with a national action program. This program consisted of two main components. The first is to spend an additional 1 billion guilders (NLG) per year to reduce class size in the first three grades of primary education. The second is the introduction of a new tax law, which gives firms the opportunity of additional tax deductions for their training expenditures. Both components got broad support in Dutch parliament as well as from unions and employers' organizations.

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

Not all training costs are deductible. First of all, they 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 forgone are excluded.

A numerical example illustrates the working of the new tax law. Consider a firm that spent NLG 200,000 on training expenditures during a fiscal year. Half of these expenditures concern older workers. The firm's total tax deduction is now equal to NLG 200,000 (the total training costs) + NLG 40,000 (20% of the total training costs) + NLG 40,000 (40% of the costs of training older workers) + NLG 12,000 (20% of the first

[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 may be.

NLG 60,000 since the firm's total training expenditures do not exceed NLG 250,000).

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

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

$$C(t) = \begin{cases} (0.51 - 0.14 \times 1_{\{\mathrm{Age} \geq 40\}}) \times C^*(t) \\ \quad \text{if } C^*(t) < 60{,}000, \\ (0.58 - 0.14 \times 1_{\{\mathrm{Age} \geq 40\}}) \times C^*(t) - 4{,}200 \\ \quad \text{if } 60{,}000 \leq C^*(t) < 250{,}000, \\ (0.58 - 0.14 \times 1_{\{\mathrm{Age} \geq 40\}}) \times C^*(t) \\ \quad \text{if } C^*(t) \geq 250{,}000. \end{cases}$$

Here $1_{\{X\}}$ is an indicator variable taking on the value of one 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%.

## III. Empirical Framework

### A. Regression-Discontinuity Design

As stated in the introduction, this article 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, we use the fact that the new tax law creates a discontinuity at the age of 40. In the first case, being age 40 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 age 40 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 of 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, Todd, and van der Klaauw (2001), who discuss identification issues, but the notation adapted is specific to the application in this article.

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 $\bar{a}$ (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 age 40 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

$$d_i = \begin{cases} 1 & \text{if } a_i \geq \bar{a} = 40, \\ 0 & \text{otherwise.} \end{cases}$$

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 because of the extra tax deduction, the (common) treatment effect.

If there is no reason to believe that persons close to $\bar{a}$ 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 $t^+ \equiv \lim_{a\downarrow\bar{a}} E(t|a)$ and $t^- \equiv \lim_{a\uparrow\bar{a}} E(t|a)$. The major identifying assumption is that there are no other discontinuities around $\bar{a}$:

Assumption RD. $E(t_{0i}|a_i)$ is continuous in $a_i$ at $\bar{a}$.

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 nontreated 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 with a fuzzy design. In a fuzzy design, assignment to treatment is not deterministic but probabilistic because it may depend as well on unobserved factors and genuine randomness. The probability of treatment therefore becomes a function of $a_i$ with a discontinuity at $\bar{a}$:

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

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 resulting from training. It can be shown (again under the assumption of a common treatment effect) that $\gamma$ can be identified by

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

where $w^+ \equiv \lim_{a\downarrow\bar{a}} E(w|a)$ and $w^- \equiv \lim_{a\uparrow\bar{a}} E(w|a)$. Because of the discontinuity at $\bar{a}$, the denominator does not equal zero. This formula is a local version of the Wald estimator and shows that RD is an instrumental variables (IV) estimator (this was first noted by van der Klaauw [2002]; see also Angrist and Lavy [1999]). Note that RD, contrary to IV, does not require that the outcome of interest ($w_i$) 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 equation (2).

### B. 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 other words, it is assumed that external treatment effects play no role.[2] In many applications this assumption can be disputed. An example concerns the estimation of returns to schooling using increases in the compulsory school leaving age as an instrumental variable (Harmon and Walker 1995). Those who are just too old to be included in the treatment group belong—as an external effect of the increase in compulsory school leaving age—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 without the change in the compulsory school leaving age. 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 those treated. Another example mentioned by Philipson is spillovers 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 spillovers, 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.[3]

[2] These external treatments effects are also referred to as spillovers or general equilibrium effects.

[3] 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.

The age-dependent tax deduction studied in this article is another example of a treatment that might have effects on the (quasi-)control group. The reduction of training costs of older workers makes it relatively more expensive to train workers younger than age 40. This might induce two different reactions from firms. The first is substitution. An employer that wants to train a worker in order to acquire some new skills or knowledge and that would send, in the absence of the treatment, a young worker to training may now decide to send an older worker to training. This implies an 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 years old, 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 age 40) out of steady state. Furthermore, while the substitution mechanism may lead to an increase of the training rates of all workers older than age 40, the postponement mechanism leads only to an increase of the training rate of workers who just turned 40 years old.

### C. Counterfactual Assumptions

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 because of the negative external treatment effect or a net increase of training for the older group. The external treatment 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 $\pi$ equals unity if the policy is in place and zero otherwise. 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 age-dependent tax deduction: $E_{99}(t|a<40, \pi = 0)$ and $E_{99}(t|a>40, \pi = 0)$. For this purpose we use information on training participation by age from a period before the introduction of the new tax law. We use data for 1994. For this year data are available that contain comparable information about the relation between training participation and age. Such data are not available for other years. Moreover, in 1994, the policy intervention was not in place, nor were there any plans for it, thereby ruling out possible anticipation effects.

Between 1994 and 1999, the overall training rate increased from 0.34 to 0.46. It is likely that this increase (at least) partially results from the other two components of the new tax law: the general deduction and the deduction aimed at small firms. As these two components may also influence the age-training profile, we are not certain whether changes in the profile between 1994 and 1999 are only due to the age-dependent component or also to these other components. A general tax deduction lowers the costs of training for the firm, which will make it worthwhile to provide more training. This probably flattens the slope of the age-training profile. But since this is not certain, we construct three different counterfactuals that accommodate a flatter profile, a steeper profile, and a parallel shift of the profile, respectively. How the deduction aimed at small firms affects the age-training profile depends on how workers of different ages are divided over firms of different size. In our 1999 data, we find no relation between firm size (measured in five level dummies) and workers' age ($p = 0.2651$). Consequently, this component has—from our perspective—the same impact as the general deduction and therefore does not require separate attention.

It is important to note that as long as assumption RD is not violated, equations (1) and (2) are both unaffected. This implies that while the other two deductions may change the age-training profile, this does not affect either our identification of the local treatment effect or the wage effect of training. The reason is that no feature of the other two deductions causes a discontinuity in the age training profile at age 40.

The next three assumptions resemble a flatter age-training profile, a steeper age-training profile, and a parallel shift of the profile, respectively.

Assumption 1a. $E_{99}(t|a, \pi = 0) = c$.

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

Assumption 1c. $E_{99}(t|a, \pi = 0) = E_{94}(t|a, \pi = 0) + c''$.

Assumption 1a implies that without the age-dependent tax deduction but with the general tax deduction, the age-training profile would be completely flat. Although this is an extreme case (since it is more likely that the profile remains downward sloping), there is not much reason to consider other cases where the profile becomes less steep. The reason for this is that our inferences with regard to external treatment effects are fairly similar for all three assumptions. Assumption 1b reflects a proportional shift of the age-training profile from 1994 to 1999. Such a shift results in a more pronounced downward slope. Finally, assumption 1c implies a parallel shift of the age-training profile between 1994 (the base year) and 1999. Combining equation (3) and assumption 1c, it now follows that:

$$\text{effect}_{<40} \equiv E_{99}(t|a < 40, \pi = 1) - E_{94}(t|a < 40, \pi = 0) - c''.$$

Similarly, equation (4) now becomes:

$$\text{effect}_{>40} \equiv E_{99}(t|a > 40, \pi = 1) - E_{94}(t|a > 40, \pi = 0) - c''.$$

Comparable expressions follow when combining equations (3) and (4) with assumptions 1a or 1b. The "only" thing needed to identify $\text{effect}_{<40}$ and $\text{effect}_{>40}$ is an estimate of the parameters $c$, $c'$, and $c''$. To identify these we need a control group that will be unaffected by the age-dependent tax deduction. The best candidates for this are young workers, and the estimations will use the assumption that workers ages 25–30 are not affected by the age-dependent tax deduction. This assumes that there are no external treatment effects for this age group. Considerations supporting this assumption are, 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.[4] Assumptions 1a, 1b, and 1c are therefore augmented with the following assumptions:

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

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

Assumption 2c. $\hat{c}'' = \bar{t}^{99}_{25-30} - \bar{t}^{94}_{25-30}$, where $\bar{t}^{94}_{25-30}$ and $\bar{t}^{99}_{25-30}$ are the average training rates for 25–30-year-olds in 1994 and 1999, respectively.

Estimators for the effects on the workers above and below age 40 can now be derived. To illustrate the procedure under the three alternative assumptions, they are written out for workers younger than age 40 (changing the subscript to > 40 gives the other estimators). Under counterfactual assumptions 1a and 2a, the following estimate of the effect on workers younger than age 40 results:[5]

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

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

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

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

[5] The appendix gives the expressions for the variances of eqq. (6) and (7).

And under assumptions 1c and 2c, the estimator becomes:

$$\widehat{\text{effect}}''_{<40} = \bar{t}^{99}_{<40} - \bar{t}^{94}_{<40} - \hat{c}''. \tag{7}$$

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

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

Comparing $\beta$ to $\beta'$ can serve as a test of the assumption that there are no other discontinuities at age 40 that affect training participation.

## IV. Data

This article uses cross-section data from 1994 and 1999. The analysis rests mainly on the 1999 data, which were collected from October 1999 to December 1999. Interviews were held by telephone using computer-aided techniques. Using proper sample weights, the data are a representative sample of the Dutch population ages 16–64. The employed persons were asked questions concerning their employment characteristics and wages, and they also responded to a set of questions addressing the training activities they undertook in the 12 months prior to the interview. 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 of 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. Although not relevant for our data, misreporting by employers would lead to an overestimate of the effect of the tax deduction on training participation. Employers have an incentive to classify workers younger than age 40 who received training as being age 40 or older.

The analysis here focuses on employed persons ages 25–55. The age bracket is narrowed to avoid complications concerning education and retirement decisions. Restricting the sample in 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 age 55. The data contain information about the worker's age at the moment of the interview. This implies that workers 40 years of age pose a problem for the analysis since it is not known whether they missed a training opportunity when they were 39 years old. They are therefore excluded from the analyses.

To apply an RD design to evaluate the age-dependent tax deduction, one needs to compare observations just below the threshold of age 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 brackets of 35–45 years of age and 38–42 years of age. These intervals cover five respectively 2 years at each side of the point of discontinuity. The observations in the ages 35–45 bracket are used to create five so-called discontinuity samples. Discontinuity sample $\pm$ 1 (abbreviated as DS $\pm$1) consists of the treatment group of 41-year-olds and the control group of 39-year-olds; DS $\pm$2 includes the treatment group of 41- and 42-year-olds and the control group of 38- and 39-year-olds. We define DS $\pm$3, DS $\pm$4, and DS $\pm$5 in a similar manner.

Appendix tables A1 and A2 report some descriptive information about 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 the age of 40 but that it is necessary to control for them as the window increases.

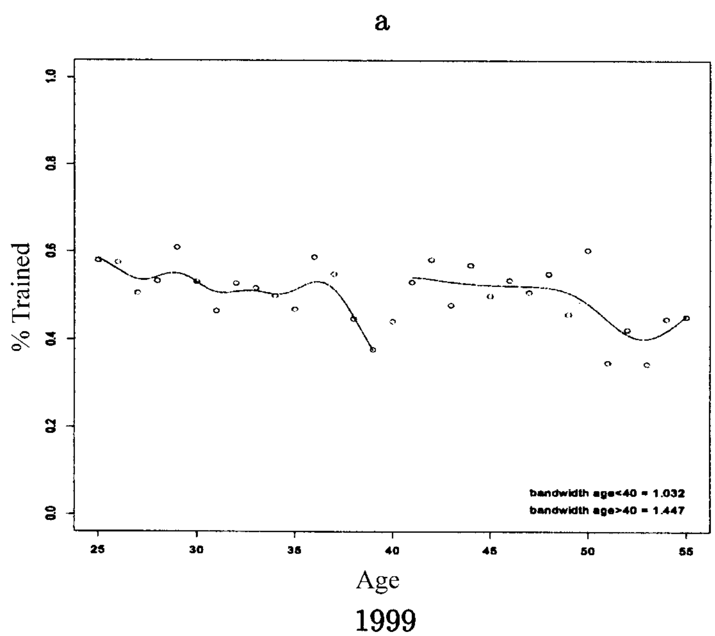
The 1994 data are needed to obtain estimates for $c$, $c'$, $c''$, and $E(t|a, \pi = 0)$ defined in assumptions 1a–1c. These data were collected as part of the Dutch wave of the so-called International Adult Literacy Survey (see Leuven and Oosterbeek [1999] for details). The important advantage of this data set is that the phrasing of the training question in this survey is identical to the phrasing used in 1999. Barron, Berger, and Black (1997) argue that existing measures of training show considerable differences and that this might result from, among other factors, the survey instruments.

## V. Evaluation of the Effects of the Tax Deduction

### A. 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. (2001), this is done by stratifying the sample on whether individuals were younger or older than age 40 before doing the local linear regression.[6] Figure 1 plots, separately for 1994 and 1999, age against the estimated training participation rate. The small circles in the figure show the actual average training rates by age. For 1999, the figure reveals a jump of 17% in these rates at the age of 40. The graph for 1994 shows

[6] The bandwidths are calculated using the plug-in approach of Ruppert, Sheather, and Wand (1995).



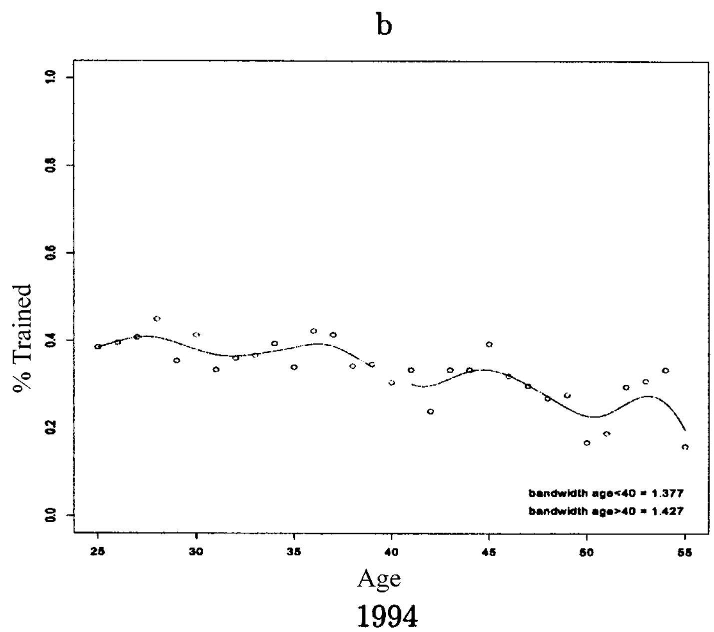




FIG. 1.—Training incidence in 1999 (*a*) and 1994 (*b*)

**Table 1**
**Training Participation Rates Discontinuity Samples (DS), 1999**

| DS | *N* (1) | < 40 (2) | > 40 (3) | Δ = (3) − (2) (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| Men and women: | | | | | | | |
| ±1 | 240 | .377 | .547 | .170 | .163 | | |
| | | | | (.069)** | (.072)** | | |
| ±2 | 385 | .425 | .555 | .130 | .120 | .241 | .273 |
| | | | | (.051)** | (.051)** | (.157) | (.155) |
| ±3 | 538 | .462 | .536 | .074 | .066 | .295 | .308 |
| | | | | (.043)* | (.043) | (.109)*** | (.108)*** |
| ±4 | 683 | .492 | .543 | .051 | .039 | .237 | .254 |
| | | | | (.038) | (.039) | (.088)*** | (.088)*** |
| ±5 | 840 | .486 | .535 | .048 | .013 | .164 | .177 |
| | | | | (.035) | (.035) | (.077)** | (.075)** |
| Men: | | | | | | | |
| ±1 | 107 | .403 | .590 | .188 | .190 | | |
| | | | | (.096)* | (.101)* | | |
| ±2 | 185 | .493 | .570 | .077 | .087 | .434 | .518 |
| | | | | (.074) | (.077) | (.224)* | (.227)** |
| ±3 | 260 | .516 | .577 | .061 | .071 | .241 | .287 |
| | | | | (.062) | (.064) | (.153) | (.154)* |
| ±4 | 336 | .554 | .565 | .011 | .013 | .248 | .269 |
| | | | | (.054) | (.057) | (.123)** | (.125)** |
| ±5 | 423 | .532 | .553 | .021 | −.005 | .127 | .120 |
| | | | | (.049) | (.051) | (.108) | (.108) |
| Women: | | | | | | | |
| ±1 | 97 | .341 | .470 | .130 | .138 | | |
| | | | | (.100) | (.103) | | |
| ±2 | 200 | .342 | .533 | .191 | .206 | −.009 | −.012 |
| | | | | (.070)*** | (.069)*** | (.220) | (.213) |
| ±3 | 278 | .397 | .472 | .075 | .067 | .383 | .442 |
| | | | | (.060) | (.059) | (.155)** | (.152)*** |
| ±4 | 347 | .418 | .506 | .088 | .074 | .205 | .266 |
| | | | | (.054) | (.053) | (.127) | (.124)** |
| ±5 | 417 | .430 | .503 | .073 | .044 | .193 | .228 |
| | | | | (.049) | (.049) | (.109)* | (.105)** |
| Controls: | | | | | | | |
| Age | | | | No | No | Yes | Yes |
| Other | | | | No | Yes | No | Yes |

NOTE.—Controls denoted by "Other" are seven education dummies, five firm-size dummies, and tenure. Controls denoted by "Age" are age and age-squared. Coefficients shown in cols. 4–7 are the coefficients of a dummy in an ordinary least squares regression. Individuals the age of 40 are excluded from the estimations. Standard errors are in parentheses.
* Significant at 10% level.
** Significant at 5% level.
*** Significant at 1% level.

that no such jump exists before the introduction of the new tax policy. There appears to be a difference in behavior comparing 1994 with 1999.

Table 1 provides estimates of the discontinuity for the five discontinuity subsamples around the age of 40. Column 1 gives the number of observations 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 2–3 of table 1 provide actual incidence rates for those younger and older than age 40, respectively. Columns 4–7 report estimates of the jump shown in figure 1*a*, with differing combinations of controls for age, age-squared, education, firm size, and tenure. These estimates are obtained from ordinary least squares (OLS) regressions with a dummy variable for training participation as the dependent variable and a dummy for age older than 40 years as an (one of the) explanatory variable(s).

The estimates in table 1, 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 2 and 3. For men and women together these estimates are all positive and decline when the range around the age of 40 is widened. The estimate that is conceptually closest to the one obtained using local linear regression is the estimate for DS$\pm$1, which is simply the difference in training rates between 41- and 39-year-olds. For men and women together this gives us a statistically significant estimate of 17%. For men this estimate is about 19% and is on the border of significance. For women the estimated effect of the training deduction equals 13% 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.[7]

Table 1, 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$\pm$1 subsample this specification cannot be estimated as it is impossible to identify three age-related coefficients with only two age levels. For all other DS subsamples, the estimates of the effect of the tax deduction go up by a factor of two or more relative to the estimates without controls for age and age-squared. This reflects what is already apparent from figure 1, namely, that—except for the intervention—training incidence declines with age. As a result, the effect of the tax deduction is in the vicinity of 25%. For the DS$\pm$2 subsample, the estimate is significant at the 10% level only, but notice that for this subsample only four age levels are used to identify the coefficients of three 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

[7] Appendix table A3 reports comparable estimates for the 1994 sample and shows that this is indeed the case.

**Table 2**
**Difference-in-Differences Estimates**

| | $\Delta_{<40}$ (1) | $\Delta_{>40}$ (2) | $\Delta = (2) - (1)$ (3) |
|---|---|---|---|
| ±1 | −.047 | .154 | .201 |
| | (.084) | (.100) | (.130) |
| ±2 | .045 | .246 | .201 |
| | (.061) | (.068)*** | (.091)** |
| ±3 | .077 | .216 | .139 |
| | (.050) | (.054)*** | (.073)* |
| ±4 | .098 | .209 | .111 |
| | (.042)** | (.047)*** | (.063)* |
| ±5 | .100 | .194 | .095 |
| | (.038)*** | (.043)*** | (.057)* |

NOTE.—Individuals the age of 40 are excluded from the estimations. Standard errors are in parentheses.
* Significant at 10% level.
** Significant at 5% level.
*** Significant at 1% level.

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 (table 1, cols. 4 and 5), while for DS±3–DS±5, the estimated effects are only significant when controls for age and age-squared are included (cols. 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% 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 table 2, column 3, are somewhat smaller than the estimates in the top panel of table 1, but clearly the effect is still very substantial, ranging from 10% for DS±5 to 20% for DS±2. Because of 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.

## B. Estimation of External Treatment Effects

The results presented so far establish that there is a substantial difference in training rates between workers below the age of 40 and those older than 40 years. This difference can be attributed to the reduction of training costs for 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 even be negative in the short run).

Table 3, column 1, provides estimates of the external treatment effect for workers younger than age 40 based on equation (5), and column 2 gives the equivalent estimate for workers older than age 40. For all five discontinuity subsamples, a significantly negative effect for workers younger than age 40 is found, ranging from 7.5% to 19%. This is a clear indication of the presence of negative external treatment effects. Without such effects, workers younger than age 40 would not be affected, and the estimated effect would equal zero. The fact that these effects are larger for the narrower discontinuity samples indicates that the negative treatment effects are mostly concentrated among workers who are just below 40 years of age. Although this shows that there are indeed external treatment effects, it does not allow us to determine whether this is caused by substitution or postponement.

For workers older than age 40, none of the estimated effects in table 3, column 2, is significantly different from zero. Since the substitution mechanism leads to an immediate increase of the training rate of older workers, these results rule out this explanation. By definition it will take some time for postponed training to catch up. The policy has only been implemented recently (January 1998), and it may take firms some time to adjust their training policy to the new rules. This implies that for workers whose training was postponed, not enough time has elapsed to observe the catching up in our sample. This explanation is corroborated by the results of table 3, column 3, which reports the total effect of the intervention. For all DS samples it is found that this sum is significantly negative. This can only occur in case of postponement. Notice that this inference does not depend on the exclusion of the 40-year-olds from the analysis. As figure 1, panel A, shows, the training rate of 40-year-olds lies between the training rates of 39- and 41-year-old workers.

Columns 4–6 and 7–9 of table 3 give alternative estimates based on equations (6) and (7), respectively. These give qualitatively fairly similar results. For younger workers the point estimates are almost the same and remain statistically significant for the narrow DS samples. For older workers the point estimates are larger except for DS$\pm$1 and remain statistically insignificant. Where these point estimates are larger they result in smaller net effects. The differences between the results for the three different counterfactuals are typically not statistically significant.

**Table 3**
**Estimates of External and Direct Effects**

| | $E_{99}(t \mid a) = c$ | | | $E_{99}(t \mid a) = E_{94}(t \mid a) + c'$ | | | $E_{99}(t \mid a) = c'' \cdot E_{94}(t \mid a)$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Discontinuity Samples | $\text{Effect}_{<40}$ (1) | $\text{Effect}_{>40}$ (2) | (1)+(2) (3) | $\text{Effect}_{<40}$ (4) | $\text{Effect}_{>40}$ (5) | (4)+(5) (6) | $\text{Effect}_{<40}$ (7) | $\text{Effect}_{>40}$ (8) | (7)+(8) (9) |
| ±1 | −.190 | −.020 | −.210 | −.232 | −.032 | −.264 | −.256 | −.040 | −.306 |
| | (.073)** | (.086) | (.113)* | (.092)** | (.104) | (.139)* | (.118)** | (.134) | (.179)* |
| ±2 | −.142 | −.012 | −.154 | −.140 | .061 | −.079 | −.142 | .093 | −.049 |
| | (.056)** | (.057) | (.080)* | (.071)* | (.075) | (.103) | (.092) | (.090) | (.129) |
| ±3 | −.105 | −.031 | −.136 | −.108 | .031 | −.077 | −.112 | .058 | −.054 |
| | (.046)** | (.047) | (.066)** | (.062)* | (.064) | (.089) | (.079) | (.077) | (.110) |
| ±4 | −.075 | −.024 | −.099 | −.087 | .024 | −.063 | −.096 | .045 | −.051 |
| | (.040)* | (.043) | (.059)* | (.056) | (.059) | (.081) | (.073) | (.072) | (.103) |
| ±5 | −.080 | −.032 | −.112 | −.086 | .009 | −.077 | −.091 | .026 | −.065 |
| | (.037)** | (.041) | (.055)** | (.053) | (.056) | (.077) | (.069) | (.069) | (.098) |

NOTE.—$t^{99}_{25-30}$ = .567 (.023), and $t^{94}_{25-30}$ = .382 (.028). This gives us estimates for $c$ = .567 (.023), $c'$ = .185 (.037), and $c''$ = 1.493 (.123). Standard errors are in parentheses.
* Significant at 10% level.
** Significant at 5% level.
*** Significant at 1% level.

## VI. Wage Effects

Previous studies have investigated wage effects of employer-provided training (see, e.g., Lillard and Tan 1992; Lynch 1992; Bishop 1994; Pischke 2001). The basic problem in estimating these effects is that selection into training is likely to be nonrandom and might depend on unobservables that correlate with wages. Comparing the wages of trained workers with wages of nontrained workers either directly or through regressing wages on training will then give a biased estimate. Two approaches may solve this problem. 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 III.A, the discontinuity around the age of 40 can be used to identify the wage returns to training by exploiting the fuzzy design of the data.

The IV estimation exploits the variation in training intensity between the treatment group and the control group caused by the age-dependent tax deduction. An important feature of this deduction is that only training for which the employer contributes at least some positive amount to the direct costs qualifies for the deduction. As a consequence we identify wage effects of training for which employers are prepared to pay. According to the standard human capital approach, such training must be specific. Otherwise the employer would not be able to recoup her investment.[8] But if the employer reaps part of the benefits of the training, then this reduces the potential wage gain for the worker.

Theory gives little guidance as to how the firm and the worker will share the returns to specific training. Hashimoto (1981) analyzes the sharing rule for a situation in which there is uncertainty about the worker's future productivity and about the value of his future outside option, in which renegotiations are not possible, and in which parties choose a fixed future wage such that it maximizes their expected joint surplus. The worker's share in the benefits of the training depends on the distribution of the worker's future productivity within the firm and the distribution of the worker's future outside option. If both distributions are the same, the parties split the benefits equally. Without further knowledge about these distributions, both very low and very high returns to specific training are consistent with the theory.

Hashimoto's model, however, also implies that the benefits of training are shared in the same way as the costs are shared. Our data sets contain

[8] Recently, Acemoglu and Pischke (1999) have shown that market imperfections may render general training effectively specific and thereby give firms an incentive to pay for this type of training as well.

some information about how parties divide the direct costs of training. It has been asked who paid these costs. The possible answers are: only the firm, only the worker, both parties, or some other party. Because this question pertains only to the direct costs and not to the (probably more sizable) opportunity costs, the information from this question does not match the relevant theoretical concepts perfectly. The frequency distributions are nevertheless suggestive. In both years it is most often the case that the firm bears the full direct costs; in 1994, this happens for almost 70% of the observed training spells, and in 1999, this has even increased to 84%. Only for 21% of the observed training spells in 1994 were the full direct costs borne by the worker; in 1999, this has decreased to 12%. The direct costs are paid by the two parties jointly or by some other party in the remaining 10% of the spells in 1994, and the remaining 4% in 1999.[9] To the extent that the pattern for the total costs is similar to the pattern for the direct costs, and to the extent that the sharing rule of the benefits equals the sharing rule of the costs, this suggests that the wage effects of training analyzed in this article are expected to be fairly small.

An interesting feature emerging from the figures just mentioned is that firms' contribution to the direct costs of training has increased between 1994 and 1999, while workers' contribution has decreased. An (albeit somewhat speculative) interpretation of this change is that the introduction of the new tax law has crowded out worker expenditures on training in favor of firm expenditures.

Table 4 reports OLS and two-stage least squares (2SLS) estimates for the various DS samples.[10] For all five 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 25–55-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 the theoretical considerations mentioned above suggest that wage returns might be low, one should note that the finding of a negligible wage return to company training participation is not uncommon. In addition to the theoretical considerations mentioned above, two factors seem to matter in this respect. First, wage returns tend to be lower in Europe than in the United States, which may be explained if market imperfections play a greater role in Europe than in the United States. Second, most U.S. studies estimate returns for younger workers, whereas this article builds

[9] The remaining sample sizes are too small to allow us to estimate the wage effects separately for workers who contributed to the cost of their training.

[10] Tables A4 and A5 report estimates for men and women separately.

**Table 4**
**Wage Returns to Training OLS and 2SLS Estimates**

| | Coefficient (1) | SE (2) | $N$ (3) | $R^2$ (4) | $F$ d40 = 0 (5) | Pr > $F$ (6) | $R^2_{d40}$ (7) |
|---|---|---|---|---|---|---|---|
| OLS: | | | | | | | |
| ±1 | .053 | (.091) | 149 | .42 | | | |
| ±2 | −.014 | (.061) | 275 | .27 | | | |
| ±3 | .001 | (.055) | 391 | .25 | | | |
| ±4 | .006 | (.046) | 503 | .25 | | | |
| ±5 | .002 | (.046) | 626 | .25 | | | |
| 25–55 | .030 | (.027) | 1,627 | .26 | | | |
| 2SLS: | | | | | | | |
| ±1 | . . . | . . . | . . . | . . . | . . . | . . . | . . . |
| ±2 | −.495 | (.654) | 275 | .13 | 3.868 | .050 | .099 |
| ±3 | −.315 | (.426) | 391 | .19 | 7.154 | .008 | .013 |
| ±4 | .107 | (.431) | 503 | .25 | 6.636 | .010 | .003 |
| ±5 | .283 | (.545) | 626 | .20 | 5.202 | .023 | .003 |
| 25–55 | −.063 | (.481) | 1,627 | .25 | 5.140 | .024 | .115 |

Note.—OLS = ordinary least squares; 2SLS = two-stage least squares. Controls include age and age-squared, seven education dummies, five firm size dummies, (job) tenure, no. of children, migrant, female, temporary, and single. Individuals who are age 40 are excluded from the estimations. Heteroscedasticity-consistent standard errors are in parentheses. They have been corrected for generated regressors.

on a sample of older workers. Examples of studies with low or zero returns are Booth (1993) for the United Kingdom and Pischke (2001) for Germany. Lynch (1992) and Veum (1995) report returns to company training incidence not statistically different from zero for U.S. National Longitudinal Survey of Youth (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 DS ±1 subsample, for example, the wage returns are identified by (see eq. [2]) the difference in wages between 41- and 39-year-old workers divided by their difference in training rates.

## VII. 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% lower for workers who are 40 years or older than for workers who are younger than age 40. This feature of the tax law is exploited to identify two effects: first, the effect of a 14% cost reduction on the probability that a worker participates in training, and second, 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 age 40 is about 15%–20% higher than the training rate of workers who are slightly younger than age 40. This difference cannot entirely be attributed to the stimulating effect of the cost reduction. On the contrary, the estimates of the external treatment effects on workers younger than age 40 indicate that these effects are so substantial that the net effect of the age-dependent tax deduction is negative. Since the data were collected less than 2 years after the implementation of the new tax law, this negative (short-run) effect is supportive of the view that the age-dependent tax deduction has led to postponement of training participation. Had the potential of spillover 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 by 15%–20%.

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 not the result of unobserved heterogeneity. These low wage returns are not unexpected. Our analysis concerns training for which the firm bears at least some of the costs. In our data set we even find that for a vast majority of the training spells, the firm pays the full direct costs. The firm will only be prepared to do so if it receives some share of the returns. This reduces the potential wage effects. It should further be noted that the wage effects identified in this article 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 article are discouraging. Our analysis does not evaluate the other two components of the new tax law (the general deduction and the deduction aimed at small firms), and we therefore have little to say about these. While the sharp increase in overall training participation rates between 1994 and 1999 from 0.34 to 0.46 is suggestive for the success of these components, there are numerous other factors that may have boosted training participation in this period. Furthermore, the absence of an effect of the age-dependent deduction also casts doubt on the effectiveness of the other two components. It is unclear why employers would be more responsive to a general tax deduction of 20% than to an age-dependent tax deduction of 40%.

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 they 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 age-dependent tax rule studied in this article, as it seems that only the costs of intramarginal training events have been deducted.

## Appendix

### Expressions for the Variances of Effect

The variance of equation (6) follows after some manipulation:

$$\begin{aligned}\mathrm{Var}\,(\widehat{\mathrm{effect}}'_{<40}) = {} & \mathrm{Var}\,(\bar{t}^{99}_{<40}) \times (\bar{t}^{94}_{<40})^2 + \mathrm{Var}\,(\hat{c}') \times (\hat{c}')^2 \\ & \times \mathrm{Var}\,(\bar{t}^{94}_{<40}) + \mathrm{Var}\,(\bar{t}^{94}_{<40}) \times \mathrm{Var}\,(\hat{c}'),\end{aligned}$$

where

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

assuming independence of $\bar{t}^{99}_{25-30}$, $\bar{t}^{94}_{25-30}$, $\bar{t}^{99}_{<40}$, and $\bar{t}^{94}_{<40}$ (see, e.g., Mood, Graybill, and Boes 1974, pp. 180–81). Similarly, it is straightforward to derive the variance of equation (7):

$$\mathrm{Var}\,(\widehat{\mathrm{effect}}''_{<40}) = \mathrm{Var}\,(\bar{t}^{99}_{<40}) + \mathrm{Var}\,(\bar{t}^{94}_{<40}) + \mathrm{Var}\,(\bar{t}^{99}_{25-30}) + \mathrm{Var}\,(\bar{t}^{94}_{25-30}),$$

also assuming that $\bar{t}^{99}_{25-30}$, $\bar{t}^{94}_{25-30}$, $\bar{t}^{99}_{<40}$, and $\bar{t}^{94}_{<40}$ are independent.

**Table A1**
**Sample Descriptives (Weighted), 1999**

| | All | | Men | | Women | |
|---|---|---|---|---|---|---|
| | Mean (1) | SD (2) | Mean (3) | SD (4) | Mean (5) | SD (6) |
| Age | 38.6 | (8.6) | 39.1 | (8.7) | 37.8 | (8.4) |
| Tenure (months): | | | | | | |
| At company | 115.7 | (107.0) | 133.4 | (115.7) | 89.7 | (86.5) |
| In function | 79.5 | (86.7) | 87.1 | (93.6) | 68.4 | (74.0) |
| No. of children | .97 | (1.13) | .95 | (1.15) | 1.00 | (1.10) |
| ln(hourly wage) | 3.375 | (.58) | 3.482 | (.56) | 3.206 | (.59) |
| ln(weekly wages) | 6.822 | (.74) | 7.084 | (.56) | 6.408 | (.80) |
| ln(hours) | 3.440 | (.42) | 3.612 | (.22) | 3.189 | (.51) |
| Single | .170 | (.38) | .189 | (.39) | .143 | (.35) |
| Migrant | .067 | (.25) | .067 | (.25) | .067 | (.25) |
| Temporary | .076 | (.27) | .056 | (.23) | .105 | (.31) |
| Female | .405 | (.49) | | | | |
| Education: | | | | | | |
| Primary | .036 | (.19) | .038 | (.19) | .033 | (.18) |
| Lower secondary: | | | | | | |
| Vocational | .113 | (.32) | .128 | (.33) | .091 | (.29) |
| General | .121 | (.33) | .094 | (.29) | .160 | (.37) |
| Upper secondary: | | | | | | |
| Vocational | .278 | (.45) | .264 | (.44) | .299 | (.46) |
| General | .063 | (.24) | .063 | (.24) | .063 | (.24) |

**Table A1 (*Continued*)**

| | All | | Men | | Women | |
|---|---|---|---|---|---|---|
| | Mean (1) | SD (2) | Mean (3) | SD (4) | Mean (5) | SD (6) |
| Tertiary: | | | | | | |
| Nonuniversity | .258 | (.44) | .264 | (.44) | .249 | (.43) |
| University | .110 | (.31) | .126 | (.33) | .087 | (.28) |
| Missing | .021 | (.14) | .022 | (.15) | .018 | (.13) |
| Firm size: | | | | | | |
| 1–10 | .106 | (.31) | .087 | (.28) | .135 | (.34) |
| 11–50 | .217 | (.41) | .205 | (.40) | .236 | (.42) |
| 51–100 | .119 | (.32) | .127 | (.33) | .107 | (.31) |
| 101–200 | .126 | (.33) | .152 | (.36) | .088 | (.28) |
| < 200 | .391 | (.49) | .405 | (.49) | .371 | (.48) |
| Do not know | .041 | (.20) | .025 | (.16) | .064 | (.24) |
| *N* | 2,326 | | 1,178 | | 1,148 | |
| *N* (weighted) | 2,326 | | 1,383 | | 943 | |

**Table A2**
***p*-Values for Equality of Treatment and Control Groups in Various Discontinuity Samples (DS)**

| | DS ±1 (1) | DS ±2 (2) | DS ±3 (3) | DS ±4 (4) | DS ±5 (5) |
|---|---|---|---|---|---|
| Tenure (months) | .307 | .073 | .008 | .000 | .000 |
| Tenure in function (months) | .983 | .259 | .023 | .000 | .000 |
| No. of children | .634 | .027 | .006 | .000 | .000 |
| ln(hourly wage) | .296 | .069 | .005 | .004 | .001 |
| ln(weekly wage) | .274 | .079 | .015 | .019 | .005 |
| ln(hours working week) | .503 | .608 | .732 | .642 | .511 |
| Single | .919 | .404 | .234 | .657 | .136 |
| Migrant | .543 | .353 | .528 | .830 | .514 |
| Temporary | .715 | .282 | .096 | .106 | .056 |
| Female | .448 | .350 | .115 | .047 | .016 |
| Education | .986 | .260 | .103 | .140 | .114 |
| Firm size | .995 | .440 | .193 | .087 | .330 |

Note.—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 in Discontinuity Samples (DS), 1994**

| | | | | Δ = (3) − (2) | | | |
|---|---|---|---|---|---|---|---|
| Discontinuity Samples | *N* (1) | < 40 (2) | > 40 (3) | (4) | (5) | (6) | (7) |
| Men and women: | | | | | | | |
| ±1 | 88 | .424 | .393 | −.031 | −.034 | | |
| | | | | (.108) | (.111) | | |
| ±2 | 168 | .38 | .310 | −.070 | −.038 | .011 | .007 |
| | | | | (.074) | (.076) | (.234) | (.234) |
| ±3 | 274 | .384 | .320 | −.064 | −.053 | −.043 | −.064 |
| | | | | (.058) | (.058) | (.157) | (.153) |

**Table A3 (*Continued*)**

| Discontinuity Samples | $N$ (1) | < 40 (2) | > 40 (3) | Δ = (3) −(2) (4) | (5) | (6) | (7) |
|---|---|---|---|---|---|---|---|
| ±4 | 387 | .394 | .334 | −.061 | −.042 | −.057 | −.044 |
| | | | | (.049) | (.048) | (.128) | (.125) |
| ±5 | 468 | .387 | .341 | −.046 | −.031 | −.08 | −.083 |
| | | | | (.045) | (.044) | (.112) | (.111) |
| Men: | | | | | | | |
| ±1 | 52 | .526 | .310 | −.216 | −.246 | | |
| | | | | (.133) | (.137)* | | |
| ±2 | 90 | .417 | .341 | −.076 | −.086 | −.527 | −.523 |
| | | | | (.098) | (.099) | (.300)* | (.297)* |
| ±3 | 148 | .449 | .333 | −.116 | −.112 | −.148 | −.206 |
| | | | | (.078) | (.079) | (.205) | (.202) |
| ±4 | 219 | .441 | .346 | −.095 | −.086 | −.169 | −.193 |
| | | | | (.067) | (.065) | (.166) | (.163) |
| ±5 | 272 | .426 | .364 | −.062 | −.056 | −.207 | −.226 |
| | | | | (.060) | (.059) | (.146) | (.144) |
| Women: | | | | | | | |
| ±1 | 36 | .184 | .563 | .379 | .345 | | |
| | | | | (.164)** | (.188)* | | |
| ±2 | 78 | .292 | .267 | −.025 | .01 | 1.078 | 1.075 |
| | | | | (.113) | (.130) | (.350)*** | (.387)*** |
| ±3 | 126 | .268 | .303 | .035 | −.003 | .187 | .254 |
| | | | | (.085) | (.089) | (.249) | (.248) |
| ±4 | 168 | .323 | .317 | −.006 | .003 | .165 | .220 |
| | | | | (.072) | (.074) | (.203) | (.205) |
| ±5 | 196 | .326 | .302 | −.024 | −.016 | .176 | .203 |
| | | | | (.066) | (.068) | (.179) | (.179) |
| Controls: | | | | | | | |
| Age | | | | No | No | Yes | Yes |
| Other | | | | No | Yes | No | Yes |

NOTE.—See table 1.
* Significant at 10% level.
** Significant at 5% level.
*** Significant at 1% level.

**Table A4**
**Wage Returns to Training OLS and IV Estimates, Men**

| | Coefficient (1) | SE (2) | $N$ (3) | $R^2$ (4) | $F$ d40 = 0 (5) | Prob > $F$ (6) | $R^2_{d40}$ (7) |
|---|---|---|---|---|---|---|---|
| OLS: | | | | | | | |
| 25–55 | .022 | (.035) | 851 | .26 | | | |
| ±1 | −.089 | (.123) | 84 | .45 | | | |
| ±2 | −.074 | (.085) | 139 | .33 | | | |
| ±3 | −.092 | (.073) | 199 | .31 | | | |
| ±4 | −.056 | (.061) | 260 | .27 | | | |
| ±5 | −.004 | (.052) | 334 | .26 | | | |
| Two-stage least squares: | | | | | | | |
| 25–55 | −.013 | (.641) | 851 | .26 | 2.491 | .115 | .223 |
| ±1 | . . . | . . . | . . . | . . . | . . . | . . . | . . . |
| ±2 | −.187 | (.335) | 139 | .32 | 10.014 | .002 | .014 |
| ±3 | −.237 | (.413) | 199 | .30 | 7.756 | .006 | .001 |
| ±4 | .164 | (.494) | 260 | .24 | 5.446 | .023 | .008 |
| ±5 | .696 | (.786) | 334 | .26 | 2.776 | .097 | .001 |

NOTE.—OLS = ordinary least squares; IV = instrumental variables. See table 4.

**Table A5**
**Wage Returns to Training OLS and IV Estimates, Women**

| | Coefficient (1) | SE (2) | $N$ (3) | $R^2$ (4) | $F$ d40 = 0 (5) | Prob > $F$ (6) | $R^2_{d40}$ (7) |
|---|---|---|---|---|---|---|---|
| OLS: | | | | | | | |
| 25–55 | .037 | (.042) | 776 | .20 | | | |
| ±1 | .093 | (.116) | 65 | .62 | | | |
| ±2 | −.029 | (.101) | 136 | .28 | | | |
| ±3 | .003 | (.088) | 192 | .25 | | | |
| ±4 | −.032 | (.077) | 243 | .26 | | | |
| ±5 | −.056 | (.083) | 292 | .26 | | | |
| Two-stage least squares: | | | | | | | |
| 25–55 | −.092 | (.726) | 776 | .19 | 2.772 | .096 | .002 |
| ±1 | . . . | . . . | . . . | . . . | . . . | . . . | . . . |
| ±2 | 3.004 | (6.206) | 136 | .29 | .288 | .592 | .140 |
| ±3 | −.396 | (.601) | 192 | .17 | 3.022 | .084 | .004 |
| ±4 | .063 | (.647) | 243 | .25 | 2.197 | .140 | .018 |
| ±5 | −.011 | (.890) | 292 | .26 | 1.465 | .227 | .015 |

NOTE.—OLS = ordinary least squares; IV = instrumental variables. See table 4.

**Table A6**
**Sensitivity of the Estimates of $c$, $c'$, and $c''$ for the Choice of Age Brackets**

| | $\hat{c}$ (1) | SE (2) | $\hat{c}'$ (3) | SE (4) | $\hat{c}''$ (5) | SE (6) |
|---|---|---|---|---|---|---|
| 25–25 | .585 | (.060) | .193 | (.089) | 1.533 | (.291) |
| 25–26 | .597 | (.041) | .225 | (.064) | 1.632 | (.233) |
| 25–27 | .566 | (.033) | .191 | (.052) | 1.528 | (.181) |
| 25–28 | .563 | (.028) | .164 | (.044) | 1.419 | (.137) |
| 25–29 | .575 | (.025) | .191 | (.040) | 1.505 | (.134) |
| 25–30 | .567 | (.023) | .185 | (.037) | 1.493 | (.123) |
| 25–31 | .555 | (.021) | .177 | (.034) | 1.475 | (.114) |
| 25–32 | .557 | (.019) | .185 | (.032) | 1.503 | (.109) |
| 25–33 | .554 | (.018) | .180 | (.030) | 1.488 | (.101) |
| 25–34 | .549 | (.017) | .173 | (.028) | 1.463 | (.093) |
| 25–35 | .543 | (.016) | .168 | (.027) | 1.453 | (.089) |

## References

Acemoglu, Daron, and Jörn-Steffen Pischke. 1999. The structure of wages and investment in general training. *Journal of Political Economy* 107 (June): 539–72.

Angrist, Joshua D., and Victor Lavy. 1999. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics* 114 (May): 533–75.

Barron, John M., Mark C. Berger, and Dan A. Black. 1997. How well do we measure training? *Journal of Labor Economics* 15 (July): 507–28.

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

Bishop, John H. 1994. The impact of previous training on productivity and wages. In *Training and the private sector: International comparisons,* ed. Lisa Lynch. NBER Comparative Labor Markets Series. Chicago: University of Chicago Press.

Booth, Alison L. 1993. Private sector training and graduate earnings. *Review of Economics and Statistics* 75 (February): 164–70.

Campbell, Donald T. 1969. Reforms as experiments. *American Psychologist* 24 (April): 409–29.

Cleveland, William S. 1979. Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association* 74 (December): 829–36.

Hahn, JinYong, Petra Todd, and Wilbert van der Klaauw. 2001. Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69 (January): 201–9.

Harmon, Colm, and Ian Walker. 1995. Estimates of the economic return to schooling for the United Kingdom. *American Economic Review* 85 (December): 1278–86.

Hashimoto, Masanori. 1981. Firm-specific human capital as a shared investment. *American Economic Review* 71 (June): 475–82.

Holzer, Harry J., Richard N. Block, Marcus Cheatham, and Jack H. Knott. 1993. Are training subsidies for firms effective? The Michigan experience. *Industrial and Labor Relations Review* 46 (July): 625–36.

Leuven, Edwin, and Hessel Oosterbeek. 1999. Demand and supply of work-related training: Evidence from four countries. In *Research in labor economics,* vol. 18, ed. Solomon W. Polachek and John Robst, pp. 303–30. Greenwich, CT: JAI.

Lillard, Lee A., and Hong W. Tan. 1992. Private sector training: Who gets it and what are its effects? In *Research in labor economics,* vol. 13, ed. R. G. Ehrenberg, pp. 1–62. Greenwich, CT: JAI.

Loewenstein, Mark A., and James R. Spletzer. 1999. Formal and informal training: Evidence from the NLSY. In *Research in labor economics,* vol. 18, ed. Solomon W. Polachek and John Robst, pp. 403–38. Greenwich, CT: JAI.

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

Mood, Alexander M., Franklin A. Graybill, and Duane C. Boes. 1974. *Introduction to the theory of statistics*. New York: McGraw-Hill.

Philipson, Tomas J. 2000. External treatment effects and program implementation bias. NBER Technical Working Paper Series T0250 (January), National Bureau of Economic Research, Cambridge, MA.

Pischke, Jörn-Steffen. 2001. Continuous training in Germany. *Journal of Population Economics* 14 (August): 523–48.

Royalty, Anne Beeson. 1996. The effects of job turnover on the training of men and women. *Industrial and Labor Relations Review* 49 (April): 506–21.

Ruppert, D., S. J. Sheather, and M. P. Wand. 1995. An effective bandwidth

selector for local least squares regression. *Journal of the American Statistical Association* 90 (December 1995): 1257–70.
Stevens, Margaret. 1994. A theoretical model of on-the-job training with imperfect competition. *Oxford Economic Papers* 46 (October): 537–62.
Van der Klaauw, Wilbert. 2002. Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review* 43 (November): 1249–87.
Veum, Jonathan R. 1995. Sources of training and their impact on wages. *Industrial and Labor Relations Review* 48 (July): 812–26.
Williamson, Oliver E. 1985. *The economic institutions of capitalism: Firms, markets, relational contracting*. New York: Free Press.