INVESTMENT IN HUMAN CAPITAL: DOES TAX-DEDUCTIBILITY MATTER?

Preliminary - Do not quote
Comments invited

Edwin Leuven  Hessel Oosterbeek

September 2005

Abstract

To stimulate investment in training by employees, the Dutch tax system allows a deduction of direct training expenditures from taxable income. This paper investigates to what extent the resulting cost reduction encourages training investments. Two different identification strategies are used. The first strategy takes advantage of the 2001 tax reform, which implied a substantial change in marginal tax rates. The second strategy uses the progressive structure of the income tax scheme and compares groups with taxable income just above or just below kinks. These strategies exploit different sources of exogenous variation and are based on different identifying assumptions. Nevertheless, the results point in the same direction: tax incentives increase human capital investment.

1 Introduction

Nowadays it is widely recognized that the level of human capital is an important determinant of countries’ prosperity. Initial education and work-

*We gratefully acknowledge valuable comments from Bas Jacobs, seminar participants in Mannheim, Paris, Stockholm, Trondheim, Uppsala. Both authors are affiliated with the Department of Economics, University of Amsterdam, NWO Priority Program Scholar and the Tinbergen Institute. Edwin Leuven is also affiliated with CREST-INSEE.
related training are the two main inputs of the accumulation of additional human capital. Whereas governments intervene in initial education in many different ways, the set of available policy instruments in the market of work-related training is limited. Moreover, while many studies have addressed the (in)effectiveness of government interventions with regard to initial education, little is known about the effectiveness of the few instruments governments can use to influence the market for work-related training.

Obviously, investing in work-related training is primarily a matter to be decided upon by private parties (employers and employees). Yet, various theoretical and empirical studies stress that private parties may under invest in work-related training and hence that the government should intervene to achieve the socially optimum investment level.

Possible instruments that a government can use to stimulate investments in work-related training include voucher schemes, individual learning accounts, certification of training (cf. Acemoglu and Pischke, 2000), tax credits and tax allowances for firms (cf. Leuven and Oosterbeek, 2004) and tax deductions for individuals.

The current paper focuses on this latter instrument that allows individuals to deduct training expenditures from their taxable income. Tax deduction of training expenditures is possible in for instance Germany, Italy and The Netherlands (in Italy against the lowest marginal tax rate), but not in countries like France and the UK (where it has recently been replaced by the now abandoned individual learning accounts). In countries like the United States, Canada and Australia training expenditures can be deducted as long as they are made to maintain existing skills.

The key question that we address in this paper is whether the tax rate against which people can deduct training expenditures from taxable income affect these expenditures.

For our analysis we use panel data from the tax register in the Netherlands covering the period 1996-2002. As we mentioned, in this country direct expenditures on training are tax-deductible. Furthermore, the Dutch income tax scheme is progressive. Consequently people who earn different taxable incomes pay different marginal tax rates and therefore incur different net costs of direct training expenditures. A simple comparison of training expenditures among individuals who pay different marginal income tax rates is, however, likely to produce a biased estimate. The first reason is that a
person’s marginal income tax rate is correlated with that person’s income. This makes it difficult to separate the effect of marginal tax rates from the effect of income. This complication is relevant for almost every empirical study on people’s responses to taxation (cf. Jappelli and Pistaferri 2003, 2004). The second complication is specific to applications dealing with tax deductibility related to investment costs. Not only the costs of an investment in training are subject to taxation, also the returns to these investments in the form of increased earnings are taxed.

We use two approaches each of which addresses both complications. In the first identification strategy we compare individuals around kink points of the income tax schedule due to jumps in marginal tax rates. In comparison with for instance the US income tax schedule, the Dutch income tax schedule has fairly large differences in the tax rates between adjacent tax brackets (see below). Individuals with a taxable income just below a kink point therefore pay a substantially lower marginal tax rate than individuals with taxable income just above that kink point. Consequently, with tax deductible direct costs of training, individuals with incomes just below the kink point have higher net costs of training than individuals with incomes just above that kink point. Because individuals in the first group have before training incomes only slightly below that in the second group, it is likely (and turns out to be the case) that their incomes in subsequent years are subject to the same marginal tax rate. As a result the two groups, although very similar, face different tax treatments of their training costs but their returns to training are treated the same.

The second approach exploits a change in marginal tax rates due to a tax reform that took place in the Netherlands between 2000 and 2001. This tax reform changed the relevant marginal tax rate for everybody, but the change in marginal tax rates was not the same for everybody. Assuming that returns to human capital investments do not start to materialize within one year after the investment, we have that the costs of investments made in 2000 and 2001 are treated differently whereas the returns are treated the same. A complicating factor is that people who are more responsive to taxation may experience a different actual change in marginal tax rates than people who are less responsive. To tackle this, the actual change in marginal tax rates is instrumented by the predicted change in marginal tax rate predicted on the basis of pre-reform income.
The results from the two approaches indicate that the level of the tax deductible rate has a positive impact on the probability that employees spend money on training and on the amount they spend. A 10 percentage points increase in the rate against which employees can deduct their training expenditures raises the probability that employees spend money on training by about 0.3-0.6 percentage point. Given that in any given year only around 4 percent of the taxpayers has deductible training expenditures, we consider this to be a substantial effect.

Section 2 reviews related empirical studies. No earlier study has focused on the effect of tax deductibility of goods invested in human capital on human capital accumulation. Section 3 describes the Dutch income tax system along with the reform that took place in 2001. Section 4 starts with a brief account of the role of tax deductibility of human capital investments in the theoretical literature on human capital and taxation, it then develops a simple model that illustrates the theoretical issues surrounding the empirical analysis. Section 5 presents the empirical specification and provides more information about the two identification strategies we use. It discusses in detail the assumptions that have to be satisfied in order to interpret the estimates as causal effects of the tax deductibility rate. Section 6 gives a description of the data. Section 7 presents the results for the local identification approach. Section 8 presents and discusses the results obtained from the approach that exploits the tax reform. Section 9 summarizes and concludes.

2 Related literature

At a general level the theoretical prediction of the effect of income taxation on human capital accumulation is ambiguous. The tax system affects human capital accumulation through at least four different mechanisms: the progression effect, the utilization effect, the deduction effect and the implicit subsidy effect.

The progression effect operates if labor taxation is progressive. In that case returns to human capital investments are charged at a (weakly) higher rate than the deductible costs of the investment. This reduces the net return of the investment and is therefore predicted to have a negative impact on human capital accumulation. The utilization effect results from the labor
supply response to taxation. Taxation reduces the net wage rate and if the
substitution effect of this dominates the income effect, this reduces labor
supply. Reduced labor supply in turn lowers the utilization of human capital
and thereby the return on human capital investments. Also the utilization
effect is therefore predicted to affect human capital accumulation negatively.
The deduction effect refers to the situation that not all costs of an investment
in human capital are tax deductible. The indirect costs of an investment in
human capital are the foregone earnings and are implicitly tax deductible
because the earnings that are foregone would have been subject to taxation.
Whether direct costs like expenditures on tuition fees, books and computer
materials are tax deductible depends on the tax code. If direct costs cannot
be deducted, a wedge is driven between the tax treatment of the (total) costs
and the tax treatment of the returns. This will lead to less human capital
accumulation. Finally, the implicit subsidy effect refers to the fact that
taxation of income from physical capital encourages investments in human
capital.

Empirical work on the relation between taxation and training decisions
is relatively rare. To the best of our knowledge, Rosen (1982) is the only
study which estimates the (total) effect of taxation on on-the-job training
decisions. He regresses on-the-job training on the internal rate of return to
training and on the marginal tax rate. On-the-job training equals one if
tenure in an employee’s present position falls short of the amount of time
required to become fully qualified in the job. Otherwise on-the-job training
equals zero. The pre-tax internal rate of return is obtained from Mincer
type wage functions which are education specific. The marginal tax rate
is predicted from a regression of an employee’s actual marginal tax rate on
age, number of children, education, non-labor income and region and type
of town dummies.

Rosen reports significantly positive effects for both the internal rate of
return and the marginal tax rate on on-the-job training. The positive sign
for the internal rate of return is in accordance with theoretical predictions.
He interprets the positive effect of the marginal tax rate as evidence that
the implicit subsidy effect dominates the other effects of taxation on training
which operate in the opposite direction (“the effect that dominates is the
one which give the individual an incentive to substitute human for physical
capital as a means for carrying consumption into the future”).

5
Some remarks with regard to this early study are appropriate. First, the measure of on-the-job training is somewhat unfamiliar by current standards. Differences in training status as measured by Rosen are likely to mainly reflect differences between more and less demanding jobs. Second, none of the instruments for the marginal rate seem to meet the exclusion restriction. These two points together suggest that the reported positive effect may easily pick up a reversed causality: individuals holding more demanding jobs pay a higher marginal tax rate (and are older, have more children, are more highly educated etc.).

More recently Gentry and Hubbard (2004) analyze the effects of marginal tax rates and of tax rate progressivity on job turnover using the PSID. Job turnover shares with training that it is an outcome with an investment dimension. Gentry and Hubbard estimate probit equations with job turnover as the dependent variable. The two key explanatory variables are the marginal tax rate and the convexity in the tax rate. The marginal tax rate is the predicted future tax rate based on household characteristics in year \( t \) and the tax code in year \( t + 1 \), assuming 5 percent earnings growth. This tax rate is supposed to capture the marginal incentives for effort at the current level of earnings as well as the relevant marginal tax rate for deductible expenses associated with job search. Convexity in the tax rate is constructed as the change in tax rates resulting from a predicted three-years increase in taxable income. This predicted increase is based on the actual distribution of 3-years period (positive) wage growth for twenty different education-age groups.\(^1\) Gentry and Hubbard find that both tax measures have a negative effect on the probability of job turnover. It is important to note that this is not a deductibility effect of the marginal tax rate, but rather a net effect which includes the effect of the marginal tax rate on effort (comparable to the utilization effect) which is predicted to be of opposite sign.\(^2\) The only thing that can be concluded from the analysis with regard to the deductibility effect is that it is dominated by the effort effect. Second, the paper accounts for endogeneity problems by including a rich set of

---

1. The mean value of this constructed variable equals 2.95 percentage points with a standard deviation of 3.12 and minimum and maximum values of -14.78 and 26.31. For comparison, the mean value of the marginal tax rate is 29.17 percent.

2. An increase in the marginal tax rate reduces the net costs of tax deductible job search activities but also reduces workers' net payoffs from an earnings increase associated with job turnover.
Related to the connection between taxation and human capital are also the many empirical studies investigating how labor supply responds to kinks in the tax scheme (see Blundell and MaCurdy (1999) for a review). Responsiveness to these kinks is a necessary requirement for the utilization effect to work. This literature does, however, not address the other requirement namely that a change in labor supply affects human capital decisions. More loosely related are also the studies which show that changes in the marginal tax rate affect taxable income (see: Feldstein, 1995; Gruber and Saez, 2002; Saez, 2003). These changes in taxable income may among other things result from changes in human capital investments.

Some studies have investigated the effect of tax deductibility on other types of expenditures. Examples include Reece and Zieschang (1985) and Glenday et al. (1986) who look at charitable contributions, Jappelli and Pistaferri (2003) who study the effect of tax deductibility on life insurance, and Jappelli and Pistaferri (2004) where the same authors investigate the effect of tax deductibility on home mortgage interest. A fundamental difference between these types of expenditures and training expenditures is that the last is an investment whereas the others are not. This is important because it implies that we also have to be concerned with the taxation of the returns.

Finally, Leuven and Oosterbeek (2004) report estimates of the effect of a policy in which firms (instead of workers) could deduct training expenditures from their tax. A specific feature of this policy was an extra deduction for training expenditures spent on workers age 40 or older. The effect of this age dependent deduction was evaluated using a regression discontinuity design. The results show a huge gap in training participation of workers just above 40 relative to workers just below 40. Further analysis suggests that this difference is mainly due to the postponement of training participation of younger workers and not to the stimulating effect of the extra deduction.

3 The Dutch income tax system

This section provides a brief description of the main elements of the income tax system during the period covered by our dataset (1996-2002). The income tax system was reformed in 2001. We first describe the system in
operation during the period 1996-2000, and then discuss the main changes that occurred in 2001. Section 5 discusses how we exploit elements of the system and of the reform to identify the effect of the tax-deductibility of training expenditures on investments in training.

In the Netherlands income tax is only levied by the central government. The basic structure of income taxation is as follows. Starting point is total gross income. This is the sum of incomes from various sources: labor, profit, capital and home ownership. To get from gross income to taxable income, gross income is reduced by a basic allowance, the size of which depends on household characteristics, and by deductions for expenditures on specific items. Expenditures on “study for career purposes” (i.e., training expenditures) is one category of expenditures that qualifies for a tax deduction. Others categories include home mortgage interest payments, alimony payments, charitable contributions and childcare expenditures. The amount of the deduction, and thereby the net costs of training expenditures, depends on the marginal tax rate. Taxable income is subject to a progressive tax scheme. Consequently, the net costs of a given amount of training expenditures is (weakly) lower for individuals with higher taxable income.

The main change induced by the 2001 tax reform is that different sources of income are no longer treated equally. Income generated by working and home ownership is still subject to a progressive tax scheme (with lower marginal rates). Income from capital and profits are taxed separately and subject to flat rates of 0.30 and 0.25 respectively.

Under the new system, training expenditures are deducted from gross income out of work and home ownership. Only if taxable income from these sources is below a certain threshold, the remaining amount can be deducted from the gross income generated by other sources.

Figure 1 shows the marginal tax rates and the kink points for the years covered by our dataset. For the pre-reform years (1996-2000) the rates pertain to taxable income from all sources. For the post-reform years (2001 and 2002) the rates pertain to taxable income out of work and home ownership. This means that the figure plots the marginal tax rates relevant for the tax deduction of training expenditures.

In 1996, 1997, and 1998 the tax schedule has three different marginal tax rates: a lowest level around 0.37, a second one equal to 0.50 and the top level of 0.60. From 1999 onwards the lower rate is replaced by two rates, which
are almost identical and in the vicinity of 0.35. Besides treating different income sources differently, the 2001 tax reform lowered the two higher rates of 0.50 and 0.60 rates to 0.42 and 0.52. The location of the kinks remained at fairly similar income levels. In Figure 1, taxable income is measured in nominal terms; when measured in real terms the year-to-year differences in the location of kink points almost vanish.

The focus of this paper is on the effect of the tax deduction of training expenditures on investments in training. For the identification we exploit the fact that due to differences in marginal tax rates different people face different net costs for the same gross investment in training. We assume other tax deductions (including the basic allowance) not to be affected by the training expenditures. The relevant marginal tax rate is therefore the rate applicable to taxable income before the deduction of training expenditures but after the deduction of all other expenditures and the basic allowance.

Whereas income taxation in the Netherlands is individualized, some deductions can be shifted from one partner in the household to the other. This is also true for the tax deduction on study for career purposes. Consequently, if someone deducts training expenses it is not certain that this
person and not this person’s partner undertook some training. Households that minimize their tax burden will shift the training expenses to the partner with the highest marginal tax rate. For this reason, we exclude observations from people who live in a household with someone with a higher marginal tax rate.\footnote{In the sample of individuals who have no partner with a higher marginal tax rate on average 3.2\% has a positive training tax deduction during a year (see also Section 6). In the excluded group of individuals living with a partner with a higher marginal tax rate this percentage is 0.6\%, whereas it would have been zero if households minimize their tax burden.}

4 Theoretical background

DepENDING on the assumptions made, different authors arrive at different conclusions about the effect of taxation on human capital formation. For instance, Heckman (1976) assumes that taxes are flat (thereby excluding a progression effect), that the value of leisure depends on human capital investments (no utilization effect) and that direct costs are fully tax-deductible (no deduction effect). Consequently, he finds that income taxation has a positive effect on human capital because only the implicit tax effect plays a role. Of course, as soon as one of the three effects which Heckman sets to zero, is at work, the net effect cannot be signed unambiguously without further knowledge about the exact sizes of the effects. Trostel (1993) for instance assumes that the goods invested in human capital are not tax-deductible. Together with a general equilibrium effect on factor prices (tax on interest leading to less physical capital, which in turn reduces wages and thereby the return on human capital) the non tax-deductibility leads in his simulation results to a negative effect of taxation on human capital. Trostel bases his analysis, however, on an assumed effect of the deduction effect. The contribution of the current paper is that it attempts to estimate this latter effect.

4.1 Theoretical framework

We introduce a simple two-period model to illustrate the relevant relations between (the extent of) tax deductibility and training investments. In this model an individual can decide to invest in training maximizing utility $U(C_1, C_2)$, where $C_t$ is period $t$ consumption. The amount of available
time per period is normalized to 1. The individual chooses which fraction $s$ of his time in period 1 will be spend on training. Time in period 1 can only be spend on work and training, time in period 2 is fully devoted to working (we thus abstract from labor supply decisions).

Direct training expenditures are denoted by $ps$. Where $p$ is the price of a unit of training in the market. Wages in period 1 are equal to $w$ and period 2 wages are equal to $wf(s)$, with $f(0) = 1$, $f' > 0$ and $f'' < 0$. We distinguish between four different tax rates: $\tau_1$ is the marginal tax rate for labor income below a certain exogenous income threshold ($Y_0$), $\tau_2$ is the marginal tax rate above that threshold, $\tau_r$ is the rate at which capital income is taxed, and $\tau_d$ is the tax rate relevant for the deduction of the direct training expenditures. With a progressive tax schedule $\tau_2 > \tau_1$.

If the individual decides not to invest in training, $s = 0$, utility equals:

$$U(0) = U(w(1 - \tau_1) - a, w(1 - \tau_1) + Ra) \quad (1)$$

where $R \equiv 1 + (1 - \tau_r)r$, the net rate of return to investment. If the individual invests in training, utility equals:

$$U(s) = U((1 - \tau_1)(1 - s)w - (1 - \tau_d)ps - a, (1 - \tau_1)Y_0 + (1 - \tau_2)(wf(s) - Y_0) + Ra) \quad (2)$$

This assumes that $wf(s) > Y_0$; with training income in period 2 falls in the higher tax bracket.

Optimization proceeds in two steps. First, the level of investment $s^*$ that maximizes equation (2) is determined, and next the two utility levels $U(0)$ and $U(s^*)$ are compared. The individual is predicted to choose $s = s^*$ if $U(s^*, a^*) > U(0, a^*)$, and $s = 0$ otherwise.

The first order conditions for maximum $U(s)$ equal

$$\frac{\partial U}{\partial s} = U_1 \cdot(-(1 - \tau_1)w - (1 - \tau_d)p) + U_2 \cdot(1 - \tau_2)wf' = 0 \quad (3)$$

$$\frac{\partial U}{\partial a} = -U_1 + U_2 \cdot R = 0. \quad (4)$$

where $U_t = \partial U/\partial C_t$, and combining them gives the following

$$\frac{(1 - \tau_2)wf'}{w(1 - \tau_1) + (1 - \tau_d)p} = R$$
and solving it for \( s \) gives a solution for the optimal level of training \( s^* \)

\[
s^* = S(\tau_1, \tau_2, \tau_d, p, R, w)
\]  

where the functional form of \( S(\cdot) \) is determined by the production function of human capital, but independent of preferences. If \( \tau_1 = \tau_2 = \tau_d \) then the optimal \( s \) does not depend on the rate at which wages are taxed (e.g. Boskin, 1975; Eaton and Rosen, 1981).

Applying the implicit function rule gives that:

\[
\frac{\partial s^*}{\partial \tau_1} = -\frac{wR}{(1 - \tau_2)f''} > 0
\]  

(6)

\[
\frac{\partial s^*}{\partial \tau_d} = -\frac{pR}{(1 - \tau_2)wf''} > 0
\]  

(7)

\[
\frac{\partial s^*}{\partial \tau_2} = -\frac{f''}{(1 - \tau_2)f''} < 0
\]  

(8)

\[
\frac{\partial s^*}{\partial \tau_r} = -\frac{r(w(1 - \tau_1) + p(1 - \tau_d))}{(1 - \tau_2)wf''} > 0
\]  

(9)

Where the inequalities follow from the second order condition for maximum utility (\( \partial^2 U/\partial s^2 < 0 \)). Higher tax rates for i) the opportunity costs, ii) the deductible direct costs and iii) the rate at which interest income is taxed, all lower the cost of investment and therefore boost investment incentives. On the other hand, a higher tax rate on the returns reduces investment incentives.

In practice tax codes may not treat \( \tau_1, \tau_d \) and \( \tau_2 \) as different elements. If marginal tax rates increase with income and training costs are fully deductible, e.g. \( \tau_2 > \tau_1 = \tau_d \), as is the case in the Netherlands, then the relation between the optimum positive investment level \( s^{**} \) and the deductibility rate is given by the expression:

\[
\frac{\partial s^{**}}{\partial \tau_d} \equiv \frac{\partial s^{**}}{\partial \tau_1} = -\frac{(w + p)R}{(1 - \tau_2)wf''} > 0.
\]  

(10)

It is important to stress the difference between equations (7) and (10). Equation (7) is the effect on training of a change of the tax deductibility rate keeping other tax rates (\( \tau_1 \) and \( \tau_2 \)) constant. Equation (10) expresses the effect on training of a change in the tax rate applicable to all (direct and indirect) investment costs. The relation between these two parameters
is the following

\[
\frac{\partial s^{**}}{\partial \tau_d} = \frac{w + p}{p} \frac{\partial s^*}{\partial \tau_d}
\]  

(11)

This shows that if one is interested in the effect of \( \tau_d \) only (equation (7)), but estimates the joint effect of \( \tau_d \) and \( \tau_1 \) (equation (10)), the effect of interest will be overestimated by factor \((w + p)/p\).

5 Empirical implementation

5.1 Specification

The analysis aims to estimate the extent to which the rate at which individuals can deduct training expenditures affects the decision to invest in human capital. Following equation (5) we estimate linear probability models that are specified as follows.

\[
s_{it} = \delta \ln(1 - \tau_{it}) + \alpha_1 E_t \ln(1 - \tau_{i,t+1}) + \alpha_2 \ln R_{it} + \lambda_i + \alpha_t + e_{it}
\]  

(12)

where \( \ln(1 - \tau_{it}) \) corresponds to \( \tau_1 = \tau_d \), and \( E_t \ln(1 - \tau_{i,t+1}) \) to \( \tau_2 \) in the model above. In the analysis \( s_{it} \) equals 1 if individual \( i \) deducted a positive amount for training expenditures from his income taxes, \( s_{it} \) equals 0 otherwise. We are interested in estimating (12) and recover unbiased estimates of \( \delta \). The main challenge for the identification of these tax deductibility effects on on training decisions is to find a source of variation in \( \ln(1 - \tau_{it}) \) (the marginal tax rate that affects the tax deductibility of training expenditures) that does not at the same time correlate with the other left hand side terms in equation (12). For the interpretation of our parameters as tax deductibility effects it is important to control for these left hand terms.

5.2 Identification

The first strategy is based on a comparison of individuals with taxable income (before the deduction of training expenditures) is either just below or just above a kink point in the tax schedule. Subsection 5.2.1 elaborates on this approach. The second method exploits the changes induced by the 2001 tax reform. Subsection 5.2.2 discusses the details of this method.
5.2.1 Local identification approach  Our first approach compares individuals with taxable income levels just above and just below kink points in the income tax schedule which are due to jumps in marginal tax rates. This corresponds to a so-called sharp regression discontinuity design (cf. Angrist and Lavy 1999; Hahn et al. 2001; Leuven and Oosterbeek 2004). Around a kink point between two tax brackets, individuals with very similar incomes face different marginal tax rates. For example, an individual who has taxable income equal to 45,000 guilders in 1996 has a marginal tax rate of 0.375, whereas someone with taxable income equal to 47,000 guilders has a marginal tax rate of 0.50 (cf. 1). Consequently, if both persons make a (gross) training investment worth 1000 guilders, the first person pays a (net) price of 625 guilders and the second person pays a (net) price of 500 guilders. Hence, this discontinuity in the tax scheme causes a difference in the net costs of training of 25 percent between two persons whose taxable income only differ 4 percent.

Comparing training expenditures between individuals with taxable incomes just above and just below the kinks in the tax schedule, may thus inform us about the effect of the tax-deductibility of training expenditures on investments in training. This requires, however, that three assumptions are fulfilled.

Define an indicator variable as follows

\[ d^k_{it}(\mu) = \begin{cases} 1 & y_{it} \in [y^{kt}, (1 + \mu)y^{kt}) \\ 0 & y_{it} \in ((1 - \mu)y^{kt}, y^{kt}) \end{cases} \]

so that \( d^k_{it}(\mu) \) equals 1 if individual \( i \) is at most \( \delta \) percent above tax kink \( k \) in year \( t \), and 0 if the individuals is at most \( \mu \) percent below the kink.

The first is that the difference in tax-deductibility is not mirrored by an off-setting difference in the taxation of the returns to training.

\[ E_t[\ln(1 - \tau_{i,t+1})|d^k_{it}(\mu) = 0] = E_t[\ln(1 - \tau_{i,t+1})|d^k_{it}(\mu) = 1] \]

In terms of the previous example: Suppose that the returns to training for the person who pays the higher price of 625 guilders are completely taxed against a rate of 0.375, and if the returns to training for the person paying the lower price of 500 guilders are completely taxed against a rate of 0.50. In that case the difference in training expenditures between people with
taxable incomes around the kink points will not identify the separate effect of the tax-deductibility but the joint effect of the tax-deductibility and the different taxation of the returns. Given, however, that the persons with taxable incomes below the kink points have their incomes close to the persons with taxable income just above the kink points, it is very likely that these groups are confronted with the same marginal tax rates in future years. In Subsection 7 we provide evidence that this assumption is fulfilled. For example, whereas the difference in marginal tax rates between the groups that are just below and just above a kink point in 1996 equals on average 0.12, the difference in marginal tax rates between these same groups is only 0.02 in 1997 and equals 0.00 from the year 2000 onwards. The same reasoning applies to the net-rate of return to capitol income:

\[ E_t[\ln R_{it}\mid d_{it}^k(\mu) = 0] = E_t[\ln R_{it}\mid d_{it}^k(\mu) = 1] \]

We also need to assume that the groups around the kinks are not systematically different in (observed and unobserved) characteristics.

\[ E_t[\lambda_i\mid d_{it}^k(\mu) = 0] = E_t[\lambda_i\mid d_{it}^k(\mu) = 1] \]

This is the usual identifying assumption in a regression discontinuity framework. To illustrate, consider for instance the kink in the 1996 tax schedule at a taxable income of 45,325 guilders. There is no reason to assume that individuals with taxable incomes in the range of \((1 - \mu)45,325-45,325\) guilders are systematically different from people with taxable incomes in the range of \(45,325-(1 + \mu)45,325\) guilders when \(\mu\) is small. The only potential threat here is that taxable incomes are manipulated with the purpose of having a higher tax deduction of the training expenditures. One way to examine whether such behavior occurs is by investigating whether taxable incomes bunch around kink points. In section 7 we report evidence that suggests the absence of bunching around kink points.

A final assumption, often implicitly made in tax studies, is that individuals know their relevant marginal tax rate when they make training investments. We share this assumption with all other empirical (and theoretical) tax studies. We realize, however, that especially for individuals with taxable incomes close to a kink point this assumption may be a strong one. This implies that the estimate obtained from the local identification ap-
approach possibly underestimates the true effect, and therefore could represent a lower bound on the true effect.

To actually conduct the regression discontinuity analysis, we need to operationalize the meaning of the phrases “just above” and “just below”. Although there is a trade-off between on the one hand achieving that people are more similar and on the other hand obtaining a sufficient number of observations, we have an additional consideration to make if would like to estimate the effect of \( \tau_d \) keeping \( \tau_1 \) constant (i.e. equation (10)). In this case we need to compensate for the fact that the net opportunity cost \((1 - t)w\)s are higher below the kink than above the kink. We try to achieve this by choosing \( \mu \) such that

\[
\int_{(1-\mu)Y_0}^{Y_0} (1 - \tau_1) y dF(y) = \int_{Y_0}^{(1+\mu)Y_0} (1 - \tau_2) y dF(y)
\]

Assuming that income is approximately uniformly distributed around the kink this implies that

\[
\mu = \frac{(1 - \tau_2) - (1 - \tau_1)}{[(1 - \tau_1) + (1 - \tau_2)]/2}
\]

the bandwidth around the kink equals the difference in net-of-tax rates above and below the kink divided by the average net-of-tax rate. By doing this separately for all 18 quasi-experiments we guarantee that the average income difference between these groups offsets the (net) difference in opportunity cost.

5.2.2 Using the 2001 tax reform In our second identification strategy we exploit a tax reform that was enacted in 2001. The 2001 tax reform changed the marginal tax rate relevant for the deduction of training expenditures. The change in the net costs of training caused by the tax reform is not the same for all individuals. People with a large share of income from capital and profits are likely to be confronted with a larger reduction than people without such income sources. Also, as Figure1 reveals, the change in marginal tax rates depends on pre-reform taxable income. Since this approach exploits changes in marginal tax rates due to the tax reform to identify the effect of tax-deductibility of training expenditures on investments in training, we difference (12) where \( t \) is the post reform year
(2001) and \( t - 1 \) the pre-reform year (2000):

\[
\Delta s_{it} = \delta \Delta \ln(1 - \tau_{it}) + \alpha_1 \Delta E_t \ln(1 - \tau_{i,t+1}) + \alpha_2 \Delta \ln R_{it} + \Delta \alpha_t + \Delta e_{it}
\]

we can simplify this equation since interest income is subject to a flat rate after the reform. This implies that

\[
\Delta \ln R_{it} = -\ln R_{i,t-1} = -\ln(1 + (1 - \tau_{i,t-1})r_t) \approx -(1 - \tau_{i,t-1})r_t
\]

since \( R_{it} \) will be absorbed by the time effects \( \alpha'_t = \Delta \alpha_t \).

Since returns to training undertaken just before the tax reform will be subject to the same tax schedule as returns from training undertaken after the reform, we assume that \( E_t \ln(1 - \tau_{i,t+1}) = E_{t-1} \ln(1 - \tau_{i,t}) \) so that \( \Delta E_t \ln(1 - \tau_{i,t+1}) = 0 \). We will therefore estimate the following equation.

\[
\Delta s_{it} = \delta \Delta \ln(1 - \tau_{it}) + \alpha_2 \tau_{i,t-1} + \alpha'_t + \Delta e_{it}
\]

Some remaining issues concerning this specification require discussion. The first issue is that if people respond to tax incentives, it is likely that their post-reform income is partially determined by such responses. This induces a potential endogeneity problem: the change in marginal tax rates may correlate with people’s responsiveness to taxation which in turn may correlate with their training decision. To address this problem, we present results in which the actual change in marginal tax rates is instrumented by the predicted change in marginal tax rate given pre-reform income: \( \Delta \ln(1 - \tilde{\tau}_{it}) \). The predicted tax rate \( \tilde{\tau}_{it} \) is calculated based on the pre-reform income \( y_{i,t-1} \) and the post-reform tax schedule.\(^4\) The identifying assumption is thus that \( E[\Delta e_{it} \Delta \ln(1 - \tilde{\tau}_{it})] = 0 \); the predicted change in marginal tax rate is exogenous conditional on the pre-reform tax rate, the pre-reform training decision and the other included observables.

\(^4\)A practical problem for the instrumentation is that we do not have information on all separate income sources in the pre-reform years. More specifically, we know individuals’ pre-reform income out of work but do not know their pre-reform income out of home ownership (the other income source determining the taxable income from which training expenditures can be deducted). However, for the post-reform years the correlation between the actual marginal tax rate and the marginal tax rate predicted on the basis of income from work equals 0.87. The omission of data on income out of home ownership therefore does not appear to be a serious problem.
6 Data

The data used in this paper come from the Dutch tax office.\footnote{The dataset is not publicly available. Permission for use of the data was granted to us in relation to a research project commissioned by the Dutch ministry of Financial Affairs.} They are the tax files of a 1.5 percent representative sample of the Dutch population for the period 1996-2002. The data are longitudinal and track individuals (and their households) from year to year.

Our two identification methods exploit variation in marginal tax rates. Because marginal tax rates are related to income and we do not want our analyses to pick up variation in income, we will condition in some of the specifications for smooth functions of income. In order not to ask too much flexibility from these smooth functions, we exclude beforehand from our analyses, observations with exceptionally low (below 5000 guilders) or high (above 150,000) incomes.

The dataset is rich on tax information. For the Netherlands it is the only source with reliable information at an individual level of taxable income and various tax deductions including the amount of any deduction related to training expenditures. We capture the information about the deduction for training expenditures into two variables: a binary variable taking the value one if training expenditures have been deducted and zero otherwise, and a continuous variable measuring the exact amount. Results will be presented for both outcomes separately.

The dataset is very limited in terms of relevant background information. It does contain information about gender, age, marital status, size of the household and economic activity, but has no information on for instance level of formal education, occupation or firm characteristics. For confidentiality reasons it is also not permitted (or possible for researchers) to merge the data from the tax register to other data sources. Moreover, the data provide no details about the training or courses related to the tax deduction.

Table 1 gives the mean values for the main variables by year for the sample remaining after applying the rules explained above. In all years the average marginal tax rate is somewhat over 0.40. Interestingly, the tax reform of 2001 which lowered the top rates by 0.08, reduced the average marginal tax rate by almost 0.02. The fraction of people with a positive tax deduction for training expenditures is slightly below 0.04 during the period
Table 1: Descriptives by year

<table>
<thead>
<tr>
<th>Year</th>
<th>N</th>
<th>marginal rate</th>
<th>fraction with positive deduction</th>
<th>deduction if &gt;0</th>
<th>Age</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>30,365</td>
<td>0.432</td>
<td>0.038</td>
<td>2277</td>
<td>39.4</td>
<td>0.22</td>
</tr>
<tr>
<td>1997</td>
<td>32,149</td>
<td>0.430</td>
<td>0.038</td>
<td>2363</td>
<td>39.5</td>
<td>0.24</td>
</tr>
<tr>
<td>1998</td>
<td>33,644</td>
<td>0.410</td>
<td>0.037</td>
<td>2390</td>
<td>39.7</td>
<td>0.26</td>
</tr>
<tr>
<td>1999</td>
<td>32,519</td>
<td>0.419</td>
<td>0.038</td>
<td>2327</td>
<td>40.1</td>
<td>0.23</td>
</tr>
<tr>
<td>2000</td>
<td>32,771</td>
<td>0.425</td>
<td>0.037</td>
<td>2359</td>
<td>40.4</td>
<td>0.24</td>
</tr>
<tr>
<td>2001</td>
<td>32,218</td>
<td>0.406</td>
<td>0.033</td>
<td>2263</td>
<td>40.7</td>
<td>0.25</td>
</tr>
<tr>
<td>2002</td>
<td>27,520</td>
<td>0.406</td>
<td>0.026</td>
<td>1909</td>
<td>40.9</td>
<td>0.26</td>
</tr>
</tbody>
</table>

1996-2000, and decreased after the tax reform to under 0.03 in 2002. The differences in training rates between the pre-reform and post-reform periods are significant at the 1%-level. The absolute amount of the training tax deduction is also fairly constant over the period and drops after the reform, meaning that after the reform deductions also decreased at the intensive margin. With regard to age and gender, we observe that the average age increases over the sample period and that the fraction of women is relatively constant.

After the reform, the average marginal tax rate decreases and the tax deduction for training expenditures decreases along both margins, suggesting that a lower marginal tax rate reduces training expenditures. This alludes already to the analysis which exploits the change in tax rates due to the reform. But whereas the patterns in table 1 may also be attributed to other factors that changed between 2000 and 2001, the analysis in section 8 compares training expenditures across individuals with larger and smaller changes in their marginal tax rates between pre-reform and post-reform years.

7 Results from the local identification approach

7.1 Assumption checks

The local identification approach builds on the assumption that the difference in tax-deductibility between people with taxable incomes just below
Figure 2: Average marginal tax rates for the 1996 “just above” and “just below” groups

and just above kinks in the tax schedule, is not mirrored by an off-setting difference in the marginal tax rates that apply to the returns to training. Figure 2 shows that this assumption is indeed satisfied. In 1996 the average difference in marginal tax rates between the groups above or below the first kink in the tax schedule equals 0.12. Following the same observations during the subsequent years 1997-2002 reveals that this difference in tax rates has already dissipated the following year. We see this for both kinks and a very similar picture emerges when we take later starting years (not shown here).

A second assumption underlying the local identification approach is that the groups that are compared (i.e. the groups around the kinks), are not systematically different in (observed and unobserved) characteristics. This is the usual identifying assumption in a regression discontinuity framework. The threat here is that taxable incomes are manipulated with the purpose of having a higher tax deduction of the training expenditures. To examine whether this occurs, we investigated whether taxable incomes bunch just above kink points. We do not find any indication of bunching at any of the kinks for the various years, a finding is consistent with results reported by Saez (2002).
7.2 Results

Each kink in the tax schedule in each of the seven years included in our dataset constitutes a quasi-experiment. This gives a total of 18 quasi-experiments. Table 2 reports the effects of tax rates on training rates. The groups are defined such that the bandwidth around the kinks (in percentages) equals the difference in marginal rates above and below the kinks separately for all 18 quasi-experiments. Results are reported for three different specifications. In the first specification no controls are included, the second specification includes controls for age, age squared, gender, marital status and economic activity. The third also controls for taxable income.

The estimates from the separate quasi-experiments are too imprecise to infer anything about the sign and the magnitude of the effect. Only three of the reported effects are significantly different from zero; these are all obtained from the quasi-experiment around the second kink in 2002. This is due to the fact that the numbers of observations located close enough around the kinks is small relative to the size of the effect. Consequently, there is a need to aggregate the information from the different quasi-experiments.

The first method that can be used for aggregation is to take a weighted average of the estimates reported in Table 2. The bottom two rows of the table report such estimates. First the separate estimates have been weighted by sample fractions (number of observations of each quasi-experiment divided by total number of observations in all 18 quasi-experiments), as in Card and Sullivan (1988). The next row reports the average effects when the separate estimates have been weighted by the inverse of their variances. This minimizes the variance of the aggregate estimate and is thus the most efficient estimator.\(^6\)

The information from the 18 quasi-experiments can also be combined by pooling the data rather than the estimates. The conventional approach is to pool the data from all the discontinuity samples, and run OLS regressions with training participation as the dependent variable and the marginal tax as the key explanatory variable.\(^7\) Table 3 reports the results.

The first three columns report the results from OLS regressions. The final three columns report the results of 2SLS estimations where the tax rate

\(^6\)Appendix A describes how the standard errors of these estimates have been calculated.\(^7\)This approach is comparable to the approach adopted by Angrist and Lavy (1999) in their analysis of the effect of class size on achievement.
Table 2: Effect of tax rate on training incidence - local estimates (± Δτ)

<table>
<thead>
<tr>
<th></th>
<th>kink</th>
<th>low rate</th>
<th>high rate</th>
<th>coef</th>
<th>s.e.</th>
<th>coef</th>
<th>s.e.</th>
<th>coef</th>
<th>s.e.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>1</td>
<td>0.375</td>
<td>0.500</td>
<td>-0.011</td>
<td>(0.014)</td>
<td>-0.023</td>
<td>(0.014)</td>
<td>-0.007</td>
<td>(0.015)</td>
<td>13,536</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.500</td>
<td>0.600</td>
<td>-0.004</td>
<td>(0.030)</td>
<td>-0.011</td>
<td>(0.030)</td>
<td>-0.009</td>
<td>(0.030)</td>
<td>3,626</td>
</tr>
<tr>
<td>1997</td>
<td>1</td>
<td>0.375</td>
<td>0.500</td>
<td>-0.019</td>
<td>(0.014)</td>
<td>-0.034</td>
<td>(0.014)</td>
<td>-0.008</td>
<td>(0.015)</td>
<td>14,062</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.500</td>
<td>0.600</td>
<td>0.010</td>
<td>(0.031)</td>
<td>0.007</td>
<td>(0.031)</td>
<td>0.001</td>
<td>(0.034)</td>
<td>3,390</td>
</tr>
<tr>
<td>1998</td>
<td>1</td>
<td>0.363</td>
<td>0.500</td>
<td>-0.028</td>
<td>(0.014)</td>
<td>-0.040</td>
<td>(0.014)</td>
<td>-0.021</td>
<td>(0.014)</td>
<td>13,603</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.500</td>
<td>0.600</td>
<td>0.071</td>
<td>(0.034)</td>
<td>0.072</td>
<td>(0.034)</td>
<td>0.069</td>
<td>(0.035)</td>
<td>2,436</td>
</tr>
<tr>
<td>1999</td>
<td>1</td>
<td>0.357</td>
<td>0.370</td>
<td>-0.385</td>
<td>(0.672)</td>
<td>-0.461</td>
<td>(0.656)</td>
<td>-0.453</td>
<td>(0.677)</td>
<td>329</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.370</td>
<td>0.500</td>
<td>-0.006</td>
<td>(0.014)</td>
<td>-0.015</td>
<td>(0.014)</td>
<td>-0.006</td>
<td>(0.015)</td>
<td>13,257</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.500</td>
<td>0.600</td>
<td>0.024</td>
<td>(0.035)</td>
<td>0.020</td>
<td>(0.035)</td>
<td>0.017</td>
<td>(0.035)</td>
<td>2,673</td>
</tr>
<tr>
<td>2000</td>
<td>1</td>
<td>0.339</td>
<td>0.380</td>
<td>-0.125</td>
<td>(0.118)</td>
<td>-0.111</td>
<td>(0.118)</td>
<td>-0.074</td>
<td>(0.112)</td>
<td>994</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.380</td>
<td>0.500</td>
<td>-0.009</td>
<td>(0.015)</td>
<td>-0.016</td>
<td>(0.015)</td>
<td>-0.002</td>
<td>(0.016)</td>
<td>12,822</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.500</td>
<td>0.600</td>
<td>-0.052</td>
<td>(0.036)</td>
<td>-0.055</td>
<td>(0.036)</td>
<td>-0.044</td>
<td>(0.036)</td>
<td>2,741</td>
</tr>
<tr>
<td>2001</td>
<td>1</td>
<td>0.324</td>
<td>0.376</td>
<td>0.014</td>
<td>(0.060)</td>
<td>0.047</td>
<td>(0.062)</td>
<td>0.053</td>
<td>(0.061)</td>
<td>2,941</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.376</td>
<td>0.420</td>
<td>-0.137</td>
<td>(0.058)</td>
<td>-0.145</td>
<td>(0.058)</td>
<td>-0.150</td>
<td>(0.059)</td>
<td>6,241</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.420</td>
<td>0.520</td>
<td>-0.021</td>
<td>(0.028)</td>
<td>-0.026</td>
<td>(0.027)</td>
<td>-0.019</td>
<td>(0.028)</td>
<td>5,979</td>
</tr>
<tr>
<td>2002</td>
<td>1</td>
<td>0.324</td>
<td>0.379</td>
<td>-0.048</td>
<td>(0.048)</td>
<td>-0.037</td>
<td>(0.048)</td>
<td>-0.038</td>
<td>(0.052)</td>
<td>2,667</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0.379</td>
<td>0.420</td>
<td>-0.059</td>
<td>(0.055)</td>
<td>-0.067</td>
<td>(0.056)</td>
<td>-0.054</td>
<td>(0.056)</td>
<td>5,107</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>0.420</td>
<td>0.520</td>
<td>-0.049</td>
<td>(0.029)</td>
<td>-0.057</td>
<td>(0.029)</td>
<td>-0.032</td>
<td>(0.031)</td>
<td>4,853</td>
</tr>
</tbody>
</table>

Controls for X: no, yes
Controls for income: no, yes

A. Average (w=N)  
-0.025 (0.007) -0.032 (0.007) -0.020 (0.007) 111,257

B. Average (w=Eff.)  
-0.015 (0.006) -0.024 (0.006) -0.010 (0.006) 111,257
Table 3: Pooled OLS and 2SLS regression estimates - Regression discontinuity sample ($\pm \Delta \tau$, N=111,257)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>$\ln(1 - \tau_{it})$</td>
<td>-0.020</td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td>(0.006)**</td>
<td>(0.006)**</td>
</tr>
<tr>
<td>Age</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.000)**</td>
<td>(0.000)**</td>
</tr>
<tr>
<td>Female</td>
<td>-0.007</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.002)**</td>
<td>(0.002)**</td>
</tr>
<tr>
<td>Household Size</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td></td>
<td>(0.001)**</td>
<td>(0.001)**</td>
</tr>
<tr>
<td>$\ln($Gross Income$)$</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)**</td>
<td></td>
</tr>
</tbody>
</table>
was instrumented with the indicator variable that equals one for individuals just above a kink and zero for those just below a kink. All regressions include year dummies. The results are very similar to those in table 2, all suggest a positive effect of the tax deduction rate on training participation. Including covariates for gender, age, age squared and marital status lead to an increase of the effect estimate (middle panel). If we also include taxable income as a control variable, the estimate is somewhat attenuated. Note that the OLS estimates are much more sensitive to the inclusion of income.

We consistently find that a 10 percentage points increase in the rate against which employees can deduct their training expenditures raises the probability that employees spend money on training by about 0.25 percentage point (with an average net-of-tax rate of 0.6).

8 Results from the tax reform

The way we use the tax reform to identify the effect of the rate at which training costs can be deducted from income taxation, is illustrated with reference to Table 4. Column entries in this table are the pre-reform marginal tax rates of 2000, row entries are post-reform marginal tax rates of 2001. The fields in the table report training rates and the numbers of observations in our data set for each combination of pre-reform and predicted post-reform tax rates. This shows, for instance, that of those with a 0.50 tax rate in 2000, 3,183 experience an increase of the marginal tax rate to 0.52, whereas all others are confronted with a tax rate reduction of at least 0.08. Our key identifying assumption is that conditional on observed characteristics (including 2000 taxable income and 2000 marginal tax rate), this change in marginal tax rates is random.

All people with a given marginal tax rate (say 0.50) in 2000, know that training expenditures in that year can be deducted from income tax against that tax rate. They are also assumed to know that any returns to these expenditures will be taxed against the new tax rates (0.3235, 0.376, 0.42 or 0.52). When these same individuals decide on their training expenditures in 2001, they know that these expenditures can be deducted from income taxation against the new marginal tax rate (e.g. 0.3235, 0.376, 0.42 or 0.52) and that any returns to these expenditures will be taxed against these same rates. Consequently, between 2000 and 2001 people have experienced a change in
Table 4: Fraction with positive expenditures in 2001 by old and predicted new rates

<table>
<thead>
<tr>
<th>Predicted new rates (2001)</th>
<th>0.339</th>
<th>0.3795</th>
<th>0.5</th>
<th>0.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>.3235</td>
<td>0.010</td>
<td>0.016</td>
<td>0.027</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(4,634)</td>
<td>(1,381)</td>
<td>(112)</td>
<td>(25)</td>
</tr>
<tr>
<td>.376</td>
<td>0.015</td>
<td>0.025</td>
<td>0.036</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>(781)</td>
<td>(9,655)</td>
<td>(194)</td>
<td>(22)</td>
</tr>
<tr>
<td>.42</td>
<td>0.064</td>
<td>0.029</td>
<td>0.033</td>
<td>0.019</td>
</tr>
<tr>
<td></td>
<td>(78)</td>
<td>(8,888)</td>
<td>(7,732)</td>
<td>(54)</td>
</tr>
<tr>
<td>.52</td>
<td>0</td>
<td>0.118</td>
<td>0.034</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(8)</td>
<td>(76)</td>
<td>(3,183)</td>
<td>(715)</td>
</tr>
</tbody>
</table>

the tax rate at which training expenditures can be deducted whereas at the same time the rate against which returns are taxed remains unchanged. Differences within each column therefore arise because of differences in the rate against which training expenditures can be deducted (because of differences in the post-reform tax rates). The higher this rate, the lower the net costs of a given training investment and therefore the higher training incidence is expected to be. The table by and large confirms this expectation. The patterns within each column are not strictly monotonic, but in all cases deviations from the expected pattern are caused by relatively small numbers of observations.

Training incidence (and training expenditures) therefore varies with tax deductibility in the predicted direction. The next step is to estimate the size of these effects and their standard errors, and to investigate how robust these estimates are with respect to various changes in the specification. The first 3 columns of table 5 presents results from regressions where the change in training incidence is regressed on the change is the marginal tax rate and various sets of covariates. Note that since we restrict the sample to the pre- and post-reform year most of the variation in the rate at which individuals can deduct is caused by the reform. Endogeneity of the post
reform rate might however be a problem, and columns (4)-(6) show results from regressions where the change of the marginal tax rate is instrumented by the predicted change of the marginal tax rate predicted on the basis of pre-reform gross income as explained above.\textsuperscript{8}

The point estimates from the first-differenced OLS regressions are all negative. Adding individual characteristics in the second column results in a more negative estimate and after adding lagged income in column (3) the estimate drops further to -0.04 and is now significant at the 5 percent level. The estimates are not only sensitive to the addition of the controls, but after instrumenting the change in the net-of-tax-rate, the negative and significant effect disappears and none of the estimates in columns (4)-(6) is statistically significant nor different of each other.

It turns out that the first-differenced specification is very restrictive. It implicitly restricts the impact of the reform to be identical and opposite signed for individuals already participating in training and those previously not participating. There is however no reason to expect that the margin by which a change in the deduction rate pushes individuals into training equals the margin by which such a change pushes individuals out of training. Fixed costs for example could generate substantial persistence in training participation which is not taken into account in table 5. We therefore present separate estimates for those who i) did not train, and ii) did train just before the tax reform.

Table 6 first shows the results for the sub-population who trained in 2000. Again the first-differenced OLS regression results are all negative, but where for the whole sample the point estimates were sensitive to the inclusion of control variables this is no longer the case. None of the effects in columns (1)-(3) is however precisely estimated. Instrumenting the change in the rate at which training is subsidized leads to much higher points estimates, all highly insignificant even though there is a very strong first stage.

Table 7 shows the results for the sub-population who did not train in 2000. Note that the sample size is much larger than in table 6. The first-differenced OLS regression results are all negative. Adding additional controls does not change the points estimates which are all strongly significant. Columns (4)-(6) show the results when the change in the rate at which

\textsuperscript{8}The instrument in highly significant in the first-stage for all the results shown here; the value of the F-test statistic is never below 62.
Table 5: Tax reform results; Full sample (N=37,538)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Δ ln(1 - τ_{i,t-1})</td>
<td>-0.0174</td>
<td>-0.0259</td>
</tr>
<tr>
<td></td>
<td>(0.0160)</td>
<td>(0.0170)</td>
</tr>
<tr>
<td>τ_{i,t-1}</td>
<td>0.0239</td>
<td>0.0408</td>
</tr>
<tr>
<td></td>
<td>(0.0199)</td>
<td>(0.0231)**</td>
</tr>
<tr>
<td>Age</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.0001)*</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Female</td>
<td>0.0055</td>
<td>0.0038</td>
</tr>
<tr>
<td></td>
<td>(0.0022)**</td>
<td>(0.0024)</td>
</tr>
<tr>
<td>Household Size</td>
<td>0.0011</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>(0.0008)</td>
<td>(0.0008)</td>
</tr>
<tr>
<td>ln y_{i,t-1}</td>
<td>-0.0054</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0024)**</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.0142</td>
<td>-0.0343</td>
</tr>
<tr>
<td></td>
<td>(0.0077)*</td>
<td>(0.0110)***</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0105)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0223)</td>
</tr>
</tbody>
</table>
Table 6: Tax reform results; Previous state trained (N=1,173)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th></th>
<th></th>
<th>2SLS</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Δ ln(1 - τ&lt;sub&gt;i,t&lt;/sub&gt; - 1)</td>
<td>-0.0335</td>
<td>-0.0291</td>
<td>-0.0464</td>
<td>0.4687</td>
<td>0.4895</td>
<td>1.1679</td>
</tr>
<tr>
<td></td>
<td>(0.2278)</td>
<td>(0.2315)</td>
<td>(0.2510)</td>
<td>(0.4800)</td>
<td>(0.5189)</td>
<td>(0.9782)</td>
</tr>
<tr>
<td>τ&lt;sub&gt;i,t&lt;/sub&gt; - 1</td>
<td>0.0320</td>
<td>0.0125</td>
<td>0.0674</td>
<td>-0.4702</td>
<td>-0.5405</td>
<td>-1.6324</td>
</tr>
<tr>
<td></td>
<td>(0.3109)</td>
<td>(0.3324)</td>
<td>(0.4600)</td>
<td>(0.5202)</td>
<td>(0.5913)</td>
<td>(1.3933)</td>
</tr>
<tr>
<td>Age</td>
<td>0.0007</td>
<td>0.0007</td>
<td></td>
<td>0.0013</td>
<td>0.0016</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0018)</td>
<td>(0.0018)</td>
<td></td>
<td>(0.0019)</td>
<td>(0.0020)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.0177</td>
<td>-0.0191</td>
<td></td>
<td>-0.0263</td>
<td>-0.0204</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0347)</td>
<td>(0.0361)</td>
<td></td>
<td>(0.0356)</td>
<td>(0.0361)</td>
<td></td>
</tr>
<tr>
<td>Household Size</td>
<td>-0.0210</td>
<td>-0.0208</td>
<td></td>
<td>-0.0206</td>
<td>-0.0217</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0117)*</td>
<td>(0.0117)*</td>
<td></td>
<td>(0.0117)*</td>
<td>(0.0118)*</td>
<td></td>
</tr>
<tr>
<td>ln y&lt;sub&gt;i,t&lt;/sub&gt; - 1</td>
<td>-0.0087</td>
<td></td>
<td>0.0886</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0507)</td>
<td></td>
<td>(0.0890)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.6279</td>
<td>-0.5806</td>
<td>-0.5082</td>
<td>-0.4322</td>
<td>-0.3875</td>
<td>-0.9336</td>
</tr>
<tr>
<td></td>
<td>(0.1270)**</td>
<td>(0.1419)**</td>
<td>(0.4453)</td>
<td>(0.2061)**</td>
<td>(0.2222)*</td>
<td>(0.5387)*</td>
</tr>
</tbody>
</table>
training is subsidized is instrumented. Compared to the OLS estimates these 2SLS estimates hardly change and are all significant.

The results in Table 7 indicate that people respond to changes in the net costs of training brought about by changes in the tax rates against which training expenditures can be deducted. The estimated effect on incidence is in the vicinity of -0.06; with an average net-of-tax rate of 0.6 this implies that, on average, an increase in the tax rate at which people can deduct training expenditures of 0.10 (10 percentage points) induces an increase in the training rate of 0.01 (1 percentage points). This is a fairly substantial effect given that average training incidence is around 3.5 percent.

9 Conclusion

This paper estimates how tax deductions, and the rate at which individuals can deduct training expenditures in particular, affects the probability of training. Evidence from two independent approaches - RD and tax reform - shows that tax deductibility has a significant and positive impact on training expenditures. A 10 percentage point increase in the deduction rate raises the probability that employees spend money on training by about 0.3-1 percentage point.

References


Table 7: Tax reform results; Previous state untrained (N=36,365)

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>2SLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Δ ln(1 − τ_{i,t−1})</td>
<td>-0.0599 (0.0112)***</td>
<td>-0.0659 (0.0118)***</td>
</tr>
<tr>
<td>τ_{i,t−1}</td>
<td>0.1071 (0.0142)***</td>
<td>0.1321 (0.0164)***</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0008 (0.0001)***</td>
<td>-0.0008 (0.0001)***</td>
</tr>
<tr>
<td>Female</td>
<td>0.0002 (0.0015)</td>
<td>0.0009 (0.0016)</td>
</tr>
<tr>
<td>Household Size</td>
<td>-0.0008 (0.0005)</td>
<td>-0.0008 (0.0005)</td>
</tr>
<tr>
<td>ln y_{i,t−1}</td>
<td>0.0023 (0.0016)</td>
<td>0.0023 (0.0016)</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.0275 (0.0055)***</td>
<td>0.0010 (0.0076)</td>
</tr>
</tbody>
</table>


A Combining estimates

We aggregate our local Wald estimates into one (average) estimate

\[ \hat{b} = \sum_i w_i b_i \]

The \( w_i \) that minimize the variance of \( \hat{b} \) are

\[ w_i = \frac{\text{var}(b_i)^{-1}}{\sum_i \text{var}(b_i)^{-1}} \]

This gives the following expression for the variance of \( \hat{b} \)

\[ \text{var}(\hat{b}) = \frac{1}{\sum_i \text{var}(b_i)^{-1}} \]

Alternatively, the local estimates can be weighted with their sample frac- tions (as in Card and Sullivan (1988)).