Labour Economics 31 (2014) 14–26

Contents lists available at ScienceDirect

Labour Economics

journal homepage: www.elsevier.com/locate/labeco

Stricter employment protection and firms' incentives to sponsor training: The case of French older workers ☆

Pierre-Jean Messe [a], Bénédicte Rouland [b,*]

[a] GAINS-TEPP, University of Le Mans, France and Centre d'Études de l'Emploi, France
[b] Department of Economics and Business, Aarhus University, Denmark

HIGHLIGHTS

- We study the effects of firing taxes on firms' incentives to sponsor training.
- Effects are identified on the basis of a policy change and theoretically illustrated.
- Taxing the layoff of older workers has no impact on their training rate.
- But substantially increases the incidence of training among younger workers.
- Effects tend to be higher among low-skilled workers.

ARTICLE INFO

Article history:
Received 14 November 2012
Received in revised form 29 March 2014
Accepted 15 July 2014
Available online 24 July 2014

JEL classification:
J14
J24
J26

Keywords:
Evaluation
Firms' training incentives
Employment protection
Older workers

ABSTRACT

This paper uses a difference-in-differences approach, combined with propensity score matching, to identify the effect of older workers employment protection on French firms' incentives to sponsor training. Between 1987 and 2008, French firms laying off workers aged over 50 had to pay a tax to the unemployment insurance system, known as the Delalande tax. In 1999, the measure was subjected to a reform that increased the tax, but only for large firms. We find that this exogenous increase substantially raised firms' incentives to train workers aged 45–49 but had no impact on the training rates among workers aged over 50. From a simple model with endogenous firing and training decisions, we give a theoretical illustration of these results.

1. Introduction

In all OECD countries, employment rates fall at the end of the working life, whatever the legal retirement age. On average, the 2012 employment to population ratio of the older population (aged 55–64) was around 55%, while it amounted to 75% for the prime age group (aged 25–54). In 1970, employment rates were respectively 56% for those aged 55–64 and 70% for those aged 25–54.[1] The average effective age at which older workers withdraw from the labor force is also decreasing: men retired at age 68.4 in 1970, but now retire at age 63.9 on average.[2] Accordingly, some countries have implemented an

☆ We are grateful to the anonymous referees and the Editor, Etienne Wasmer, for valuable comments and suggestions that substantially improved the quality of this paper. Moreover, we thank Francois Langot, Hélène Turon, Francois-Charles Wolff, as well as participants at the 2011 CNRS Winter Thematic School on "Public Policies Evaluation", the 2011 Conference of Applied Microeconomics (JMA), the 2011 Journées Louis-André Gérard-Varet, the 2011 Annual Conference of the European Association of Labour Economists (EALE), the 2011 Annual Workshop of the Search and Matching Research Group (SaM), the 2011 Annual Workshop of the Centre for Research in Active Labour Market Policy Effects (CAFÉ) on the Evaluation of Active Labor Market Policies, the 2012 ADRES Doctoral Conference, the 2012 Annual Meeting of the Society of Labor Economists (SOLE), the Workshop on "Theory and Practice of Program Evaluation" in Luxembourg, and seminar participants at the University of Evry, the University of Le Mans, the University of Strasbourg, IZA, Aarhus University, the University of Edinburgh and the Ohio State University for their helpful comments and discussions. All remaining errors and shortcomings of the article are our own.

* Corresponding author.
E-mail addresses: Pierre-Jean.Messe@univ-lemans.fr (P.-J. Messe), brouland@econ.au.dk (B. Rouland).

[1] OECD data on Labour Force Statistics by sex and age, extracted from OECD data bank (http://stats.oecd.org/).

[2] OECD time series: http://www.oecd.org/els/emp/ageingandemploymentpolicies-statisticsonaverageeffectiveageofretirement.htm.

http://dx.doi.org/10.1016/j.labeco.2014.07.004
0927-5371/

age-specific employment protection. Thus, in Belgium, Finland, France, Japan, Korea and Norway, the employment protection legislation is more stringent for firms laying off workers aged over 50 (OECD, 2006). The effects on employment flows have been documented: stronger employment protection should foster long-term relationships between older workers and employers[3]. For instance, Schnalzenberger and Winter-Ebmer (2009) show that an increased tax for the layoff of older workers in Austria significantly decreased their layoff rates. A similar regulation for France (the "Delalande tax") has been analyzed by Behaghel et al. (2008). The authors show that the more stringent schedule of the tax following the 1999 reform (a change that is under consideration in this paper) led to a notable decrease in the layoffs of older workers in large firms.

In this paper, we study the effects of age-specific employment protection on firm's incentives to train the workforce. To the best of our knowledge, this is an interesting question that has not been studied yet. While lifelong learning is an important European public policy concern,[4] access to training is deeply unequal among employed workers: it increases with the number of qualifications and the size of the firm, and decreases with age, specifically for workers aged over 50 (Bassanini et al., 2007). We wonder here whether age-specific employment protection could help to reduce age disparities in access to training. More generally, in line with Picchio and van Ours (2011) who show that training may improve the employability of workers, even for older workers, our study provides new insights for policies aiming to maintain older workers in employment.

There is an extensive literature on firms' incentives to sponsor training, starting with Becker (1962). In a competitive environment, Becker conjectured that firms should only pay for specific training since it increases workers' productivity only in their current job while general training should be wholly financed by workers themselves as it improves their productivity in future jobs as well. Acemoglu's contributions in the late 1990's considering frictional markets have led to the discarding of Becker's point of view: firms do have incentives to finance general training and firms' investments in specific training are subjected to hold-up issues.[5] More specifically, the link between employment protection and firms' incentives to sponsor training has been also studied. Some contributions look at the effects on workers' investment in their own training, highlighting how employment protection stimulates their willingness to invest in firm-specific knowledge (Wasmer, 2006 and Belot et al., 2007). Other papers put the emphasis more on firms' incentives to invest in training. Acemoglu (1997), Fella (2005) and Lechthaler (2009) show that firing costs help raise firms' investment in general training, so optimal training choices can be achieved. As to investments in specific training, Chéron and Rouland (Chéron and Rouland, 2011) argue that the implementation of both employment protection and training subsidies is necessary to restore social efficiency since firms' training and firing decisions are strongly complementary. Moreover, in a worldwide study, Pierre and Scarpetta (2004) have shown that firms facing tight regulations on the labor market invest more in training to accommodate the workforce to needs created by new technologies. The contribution of our paper is to examine firms' training choices in a setting of age-specific employment protection.

Our approach and key results can be summarized as follows. First, from a theoretical point of view, a firing tax targeted on workers who reach a certain age has age-differentiated effects on firms' training decisions. In particular, we show that such a firing tax has no impact on the training rates of eligible workers for the tax (due to the imminence of retirement which makes expected marginal returns on training close to zero). However, it has a positive impact on the incidence of training among younger workers, not eligible for the tax but close to the threshold age. This result comes from a complementarity effect between firing and training decisions of firms. The idea is that the increased firing tax after age 50 changes the firing policy of firms, both for workers under and over 50. Therefore, knowing that dismissals will cost much more after age 50, firms have strong incentives to engage in training activities before the worker turns 50 to increase productivity and job value, which in turn will lower the probability of a job destruction after 50 (which would cost much more because of the tax increase). Secondly, we take advantage of a change in the Delalande tax' schedule in 1999 to properly identify its effect on training rates for workers aged over 50 and for workers close to the 50 year threshold. As this change only concerned large firms (employing 50 employees or more), we can use a difference-in-differences approach combined with propensity score matching to compare the incidence of training among workers employed in small and large firms, before and after the reform. Specifically, we find a substantial positive effect of the change on firms' incentives to provide training, but only significant for the 45–49 age group. We also show that the treatment effect appears to be greater among less productive jobs. These findings are consistent with our theoretical predictions.

The paper proceeds as follows. In Section 2, we document the institutional background of the Delalande tax and of training in France. Section 3 then gives theoretical illustrations of the role of employment protection on training incentives. Section 4 presents data and associated descriptive statistics. The empirical analysis is discussed in Section 5. Section 6 concludes.

2. Institutional background

2.1. The Delalande tax

From its introduction in 1987 until its cancelation in January 2008, French firms laying off workers aged over 50 had to pay the Delalande tax to the unemployment insurance system.[6] The amount of the tax was proportional to the worker's gross wage at the time of layoff. The 1992 reform exempted from taxation all layoffs of workers who were hired after the age of 50. The tax was only due if the worker was employed under a permanent contract and only the private sector was concerned. In 1999, the Delalande tax' schedule was increased, but the change only concerned firms with more than 50 employees (referred to as large firms hereafter). In the empirical analysis, we will use this change as an identification strategy by comparing training rates in small (below 50 employees) and large (over 50 employees) firms, before (in 1998) and after (in 1999) the reform. The rise in the tax was implemented in a context of rapidly growing employment from which all categories of workers benefited, except older unemployed workers. Table 1 shows how the amount of the tax varied after the 1999 reform.

2.2. The French training system

In 2011, out of the 32 billion euros (i.e. about 1.6% of GDP) spent on initial and continuing vocational training in France, more than 40% (13.6

[3] The general conclusion reached in the extensive literature on employment protection legislation is that employment protection measures do not have a significant impact on steady-state employment, but are likely to influence employment dynamics (see Young, 2003 for a review). More precisely, with fewer job terminations and less job creation, employment protection legislation is known to reduce inflows into unemployment and outflows from employment, while also lowering outflows from unemployment and inflows into employment. However, this indirect negative effect of employment protection on the overall employment rate does not concern older workers as their hiring rate is very low. Therefore, only the direct effect on firing rates matters regarding older workers.

[4] Lifelong learning is a core objective of the European Union. In particular, lifelong learning has a key role in the "Flexicurity" strategy. It is about developing flexible work organizations where workers can keep their skills up-to-date, which gives them a more flexible environment for changing jobs and a higher probability of re-employment.

[5] See Acemoglu and Pischke (1998) and Leuven (2005) for surveys.

[6] The threshold-age was 55 in 1987 but was lowered to 50 after the 1992 reform.

Table 1
Delalande tax schedule according to the age of the laid off worker (monthly gross wages).

		Worker's age								
		50	51	52	53	54	55	56-57	58	59
Jan. 1993–Dec. 1998	All firms	1	1	2	2	4	5	6	6	6
Since Jan. 1999	More than 50 employees	2	3	5	6	8	10	12	10	8
	Fewer than 50 employees	1	1	2	2	4	5	6	6	6

Source: Behaghel, Crépon and Sédillot (Behaghel et al., 2008), legislative texts.
Notes: For each age group, the table displays the tax due by the firm to the unemployment insurance system if it lays a worker off. The tax is a function of previous wages, and is stated in months of gross wages.

billion) represented training expenditures for the workers employed in the private sector (Delort (2013)).[7] Firms were by far the main contributor (84% of the training expenditures for the workers employed in the private sector, i.e. 11.4 billion). A distinctive feature of the French vocational training system is the existence of mandatory contributions. Specifically, besides a direct investment in the staff training plan, all private-sector companies in France have to devote a specific percentage of their total annual wage bill (at least) to training activities. Mandatory contributions vary with the firm's size. From December 2013 (the latest reform), the minimum requirement for companies employing 10 workers and more is 1% of the total wage bill, while smaller firms (1 to 9 employees) have to devote 0.55% of their total wage bill to training activities.[8] Firms not fulfilling this legal requirement have to pay a tax instead (same amount), leading to the so-called "pay-or-train" system. In practice, mandatory contributions are collected by social partners.[9]

While training expenditures are considerable in France, the training system does not work well, as pointed out in many reports.[10] There are many problems, mainly arising from the mandatory and centralized contributions. First, a "pay-or-train" system cannot be efficient since small firms have no incentive to spend more than the legal requirement, even if their employees need more training.[11] Moreover, larger firms deciding to spend more than the minimum requirement have no incentive to account for the positive externalities that training could create. Only the private return matters in the decision. Furthermore, for the sake of equity, shared contributions should be redistributed toward unskilled and low-skilled workers, but the highest educated workers are the ones who actually get trained, because of a higher private return. Another major problem of the French training system arises from the central place held by the social partners (OPCAs) in the system. While collecting mandatory contributions, OPCAs are legally authorized to use up to 10% of collected amounts to finance their management costs, as well as 1.5% to finance social partnerships. Besides these levies, OPCAs have full control over the redistribution among firms of collected contributions, as well as over the providers.[12] The lack of tight regulatory constraints and of independent certification over the quality of the training exacerbates their dominant role.

3. Theoretical illustration

3.1. Model environment

We give theoretical illustrations of the role of age-specific employment protection on firms' incentives to provide training in the following simplified environment. We consider a discrete time model with a finite life time horizon for workers à la Chéron et al. (2011), extended to account for firms' investment in training. In this setting, older workers exit the labor market at the exogenous age T, perfectly known by employers. The model has two periods i, for the two last periods of working life before retiring (denoted $T-2$ and $T-1$).

A productive unit is the association of one worker and one firm who are already matched.[13] The productivity of a worker is the sum of a random component ε and a deterministic one $y^i(k^i)$, derived from training investments in specific human capital k^i during both periods.[14] Lastly, in any period, workers earn an exogenous wage b.

The time of events and of decisions is as follows. First, at the beginning of the period, an idiosyncratic productivity shock hits jobs. A new job productivity ε is then drawn in the general distribution $G(\varepsilon)$ with $\varepsilon \in [0, \overline{\varepsilon}]$, and the firm has no choice but either to continue production or to destroy the job. Firing workers aged $T-2$ implies no specific cost, while a firing cost F has to be paid when a firm fires a worker of age $T-1$. Firms decide to terminate any job with a productivity below an (endogenous) productivity threshold denoted R^i that depends on the amount of training investments. The job destruction rate is then determined by $G(R^i)$. Secondly, firms decide on the investment in firm-specific skills k^i that determines both the human capital of workers for the period $y^i(k^i)$ and their overall human capital level.[15] There is no human capital depreciation and all skills are allowed to accumulate between the two periods. More precisely, $y^{T-2} = y^{T-2}(k^{T-2})$ and $y^{T-1} = y^{T-1}(k^{T-2}, k^{T-1})$.

3.2. Firms' decisions

3.2.1. Firing decision

For a firm, the intertemporal value of a filled job depends both on the worker's human capital $y^i(k^i)$ and on the idiosyncratic component ε. This value is denoted by J^i. We assume that firms pay all the training

[7] Training expenditures for all employed workers, from both the private and the public sector, amounted to 62% of the total expenditure on vocational training in 2011. In comparison, training young people accounted for 25% of the total expenditure while training unemployed workers represented 13% of it. See Delort (2013) for more details.

[8] Before the December reform, the minimum requirement for companies employing 20 workers and more was 1.6% of the total wage bill. Contributions for firms employing between 10 and 19 workers had to represent 1.05% of their total wage bill, while smaller firms (1 to 9 employees) had to devote 0.55% of their total wage bill to training activities.

[9] In 2011, 46% of the total expenditure of firms on training (direct investment and legal requirement) was collected by the social partners, the *organismes paritaires collecteurs agréés (OPCAs)*.

[10] See Cahuc et al. (2011) for instance.

[11] It is well-known that workers in small firms are less educated on average than workers employed in large firms. This is also the case in our data set (see Section 4 on descriptives statistics).

[12] See Labruyère (Labruyère, 2005) for a detailed description of the role of the social partners.

[13] As we only focus on training and firing decisions of firms, we do not account for the hiring process. Therefore and for the sake of simplicity, we consider associations which are already productive.

[14] The additive form of the output of the match we assumed between an endogenous component ($y^i(k^i)$) and another exogenous one (ε) clearly simplifies calculations but also fits the usual definition of training. Usually, training is considered as a way of improving workers' skills. Without training, workers are still able to produce but at lower productivity levels. To mention only a few papers using the same specification, Lechthaler (2009) and Belot et al. (2007) consider an additive form of the output of the match as well (within the framework of endogenous human capital and productivity shocks).

[15] The function $y^i(k^i)$ is supposed to be strictly increasing and concave, with $y(0) = 0$.

costs $C(k^i)$[16]. At each period i, corresponding Bellman equations satisfy:

$$J^i = y^i\left(k^i\right) + \varepsilon - b - C\left(k^i\right) + \beta\left[\int_{R^{i+1}}^{\bar{\varepsilon}} J^{i+1} dG(x) - G\left(R^{i+1}\right) F^{i+1}\right] \quad (1)$$

with $F^{T-1} = F$, $F^{T-2} = 0$ and $J^T = 0$ (terminal conditions). It is worth noticing that, at each period i, k^i is the best response of the firm given the training investment made in the previous period. In addition, productivity thresholds R^i depend on the accumulated specific human capital k^i. Integrating by parts and rearranging terms, Eq. (1) comes to:

$$J^i = y^i\left(k^i\right) + \varepsilon - b - C\left(k^i\right) + \beta\left[\int_{R^{i+1}}^{\bar{\varepsilon}} [1 - G(x)]dx - F^{i+1}\right]. \quad (2)$$

The endogenous job destruction rule leads to a reservation productivity R^{T-1} in $T-1$ defined by $J^{T-1} = -F$, as a firing cost has to be paid at this period. The reservation productivity R^{T-2} in $T-2$ is defined by $J^{T-2} = 0$, as dismissals do not imply any cost. Given the terminal conditions, we obtain:

$$R^{T-1} = b - F - \left[y^{T-1}\left(k^{T-1}\right) - C\left(k^{T-1}\right)\right]. \quad (3)$$

In addition, in $T-2$, the job destruction rule is given by the following expression:

$$R^{T-2} = b - \left[y^{T-2}\left(k^{T-2}\right) - C\left(k^{T-2}\right)\right] + \beta F - \beta\int_{R^{T-1}}^{\bar{\varepsilon}} [1 - G(x)]dx. \quad (4)$$

Therefore, as already argued by Behaghel (2007) and Chéron et al. (2011), the firing cost F has an opposite effect on R^{T-1} and on R^{T-2}. F leads to reducing the job destruction rate in $T-1$ (labor-hoarding effect), while it increases it in $T-2$ because firms anticipate the future firing tax. In addition, training investments improve job tenure by increasing future productivity gains. In particular, the higher the option value of filled jobs (i.e. the expected gains in the future, depending on the training investments), the weaker the job destructions.

3.2.2. Training investment decision

Firms choose the amount they invest in specific skills at each period in order to maximize the net expected value of a filled job. In this way, firms decide on the sum they will invest in specific training so that the expected marginal return on investment is equaled to its marginal cost. Specifically, the marginal return depends on present and future expected profits. In $T-1$, the expected value of a filled job only depends on the instantaneous profit as workers retire in T. But the expected profit in $T-1$ is also determined by the training investment carried out in $T-2$. Therefore, denoting $k^{T-1} = \kappa(k^{T-2})$,[17] the optimal investment decision rule in $T-1$ conditionally on k^{T-2} is as follows:

$$C'\left(k^{T-1}\right) = \frac{\partial y^{T-1}}{\partial k^{T-1}}. \quad (5)$$

By increasing productivity in $T-1$, the training investment in $T-2$ affects the expected job value in $T-1$. Therefore, the optimal investment decision rule in $T-2$ is defined by:

$$C'\left(k^{T-2}\right) = \frac{\partial y^{T-2}}{\partial k^{T-2}} + \beta A\left[1 - G\left(R^{T-1}\right)\right] \quad (6)$$

where $A = \frac{\partial J^{T-1}}{\partial k^{T-2}}$ gives the net marginal return in $T-1$ on the training investment carried out in $T-2$ (provided the job is not terminated).

[16] With $C'(0) = 0$, $C'(k^i) > 0$ and $C''(k^i) > 0$.
[17] With $\kappa'(k^{T-2}) \geq 0$.

3.3. The effect of the age-specific firing cost on firms' training incentives

We now investigate the theoretical effects of an increase in the age-dependent firing tax F on firms' incentives to train workers aged $T-1$ (eligible for the tax) and workers aged $T-2$ (non-eligible). First, from Eq. (5), it is clear that an increase in F will not affect the optimal training investment in $T-1$, as workers in $T-1$ are very close to retirement. Expected marginal returns on training are nil. Because the firing tax influences training decisions through expected gains of training, the propensity of firms to train older workers is not affected by a change in this tax, because of their short horizon. This result is intuitive and consistent with the existing literature on the positive link between the time before retiring and the probability of getting training (Khaskhoussi and Langot, 2008; Montizaan et al., 2010).

Then, we look at the effect of an increase in F on the training rate of the previous cohort of workers. Combining Eqs. (6) and (3) implies:

$$\Phi\left(k^{T-2}, F\right) \equiv -C'\left(k^{T-2}\right) + \frac{\partial y^{T-2}}{\partial k^{T-2}} + \beta A\left[1 - G\left(R^{T-1}\right)\right] = 0. \quad (7)$$

Differentiating Eq. (7) leads to $\frac{\partial k^{T-2}}{\partial F} = -\frac{\Phi_2\left(k^{T-2}, F\right)}{\Phi_1\left(k^{T-2}, F\right)}$ where $\Phi_1(k^{T-2}, F)$ is necessarily negative to get an interior solution. Therefore, the impact of F on k^{T-2} only depends on the sign of $\Phi_2(k^{T-2}, F)$:

$$\Phi_2\left(k^{T-2}, F\right) = \beta A\left[\frac{\partial\left[1 - G\left(R^{T-1}\right)\right]}{\partial F}\right] \equiv \beta A g\left(R^{T-1}\left(k_{T-2}, \kappa\left(k^{T-2}\right)\right)\right) > 0. \quad (8)$$

The impact of F (due in $T-1$) on k_{T-2} rests on the way the tax affects the probability that the job will not terminate after a productivity shock. Specifically, an increase in the firing cost clearly raises firms' incentives to provide workers with training in $T-2$. This comes from a complementarity effect between training and firing decisions.[18] In particular, firms have strong incentives to protect matches against bad productivity shocks when dismissals in the next period are subject to the tax. The training investment increases the worker's productivity and therefore the (intertemporal) job value, which in turn reduces the risk of layoff that would cost much more at the next period because of the tax. It is worth noticing that in this model, the effect of age-specific employment protection on firms' training incentives does not depend on the skill-level of jobs, given that we assumed that at the beginning of each period, all jobs are hit by a productivity shock. But even without shocks, as the job destruction rule changes when the worker gets older, some jobs, i.e. the less productive ones, may be not robust to aging. It implies that if the probability of a shock at the beginning of each period is lower than one, age-specific employment protection encourages firms to invest in training in $T-2$ but only for low-skilled workers. Indeed, for highly productive ones, the risk of layoff at the next period is nil even without training.

4. Data and descriptive statistics

4.1. Data description

To assess the effect of the Delalande tax reform on firm-provided training, in this study we use two complementary French databases. The first is a cross-sectional survey on continuing training entitled "Formation Continue", which is a supplementary survey to the Labour

[18] Chéron and Rouland (2011) show that job destructions and training investments are highly complementary since firms have strong incentives to invest in training to protect matches from idiosyncratic productivity shocks. Expected productivity gains due to training investments raise the job tenure, which in turn encourages firms to invest more.

Force Survey (LFS hereafter). A sample of 28,700 people under 65 years of age and not in education was interviewed in March 2000. The aim of the survey was to gather information on all the training spells undertaken by individuals, on the context in which training decisions are taken and on the constraints determining participation in training. Most training spells were recorded, regardless of purpose (explicitly work-related or personal), duration or type (courses, block-release, work-based or self-directed). Specifically, training spells are grouped into three main periods: i) from leaving school to February 1998, ii) from March 1998 to December 1998, and iii) from January 1999 to March 2000. The training survey gives information about who paid for the training session. We decide to only consider firm-financed training sessions, which represented around 80% of all training spells of those aged 45–54 employed in the private sector in 1998 or in 1999. The survey also describes the purpose of the training activity, which has to fit into one of the following categories: i) to adapt to the job, ii) to switch to another job or to get a job, iii) to obtain a diploma or a certification, iv) to execute political duties, v) personal or cultural reasons, and vi) no specific reason. In our theoretical model, training is a response of firms to some changes in the working environment that may affect workers' productivity. Therefore, we limit the analysis to training spells aimed at adapting to the job,[19] which represented more than 85% (89%) of the firm-financed training sessions in which older workers employed in the private sector participated in 1998 (1999). Actually, this reflects the fact that new technologies and organizations require continuing learning.

We merge these data with the LFS conducted in France in 1998 and in 1999.[20] This is a rotating panel since exactly one-third of the sample is dropped each year and is replaced with a new and comparable sample drawn from the current population. The size of the LFS is about 135,000 individuals who are annually interviewed about their situation in the labor market. Therefore, we have detailed information over two years for two-thirds of the original 1998 sample. For each respondent, the LFS contains detailed information about socio-demographic individual characteristics, as well as job and firm characteristics.[21]

4.2. Sample selection and descriptive statistics

We define our sample in the following way. We start by excluding farmers and the self-employed as well as individuals working in the public sector.

As on-the-job training spells are observed only for employed respondents, it may produce a selection bias. Indeed, we could guess that after the reform, large firms would fire more workers aged 45–49 in anticipation of the tax, retaining only the more productive workers. In that case, a positive effect of the reform on the training rates among this age group of workers could only arise from a selection effect. Examining some statistics from the French LFS, we see that the annual separation rates are larger in small firms than in large firms, both before and after the reform. More precisely, before the reform the separation rates among workers aged 45–49 are 3.83% and 7.71% for large and small firms, respectively. After the reform, these rates are 1.36% and 6.59%, respectively. Thus, the difference in separation rates between large and small firms is larger after the reform (−5.2%) than before the reform (−3.9%). Consequently, this negative difference-in-difference (equaled to −1.3%) does not show evidence of greater firing after the reform for the 45–49 employed in large firms. Besides, the selectivity effect, if any, would be stronger for small firms not concerned by the 1999 reform.

In addition, more than 93% of the workers aged 45–54 employed in 1998 are still in employment in 1999 while the corresponding share for workers aged 55–59 is about 80%. So the selection bias may be quite strong for workers aged 55 and over but is negligible for those aged 45–54. Therefore, we focus on the population of respondents aged 45–54 in March 1998 in order to reduce the likelihood of a selectivity effect.[22]

As the time before retiring is expected to have a strong impact on firms' training decisions (Montizaan et al., 2010), we want to control for it. But this implies restricting the analysis to men. Indeed, while unemployment episodes are included in the number of contributive periods, the proxy we use for the retirement age[23] does not account for breaks in careers due to maternity and parental leave. But the latter are undoubtedly mostly taken by women in France and could lead to miscalculating their distance to retirement, as women can go on several leaves in a row in the case of new births. Therefore, women are excluded from the analysis. Dropping the few missing values (mainly because of missing information on firms' size) and keeping only workers who were working both in 1998 and in 1999, we get a balanced panel with 955 observations for each year. We investigate the participation of these individuals in a firm-financed training spell (aimed at adapting to the job), while employed, between March 1998 and December 1998 for the pre-reform period, and between March 1999 and December 1999 for the post-reform period.[24]

Table 2 shows some descriptive statistics about our sample. First, we provide some figures about the incidence of training in small and large firms between 1998 and 1999 for the 45–49 age group and the 50–54 age group. As mentioned in Section 1, we consider firm-financed training spells that aimed at adapting to the job as baseline. We see that between 1998 and 1999, the incidence of firm-financed training in small firms varies similarly for the 45–49 and the 50–54 age groups, while in large firms it increases much more for the 45–49 age group than for the 50–54 age group, regardless of the purpose. However, this picture is not so clear for training spells not sponsored by firms. In addition, when we examine the length of the firm-financed training sessions aimed at adapting to the job, we see that training sessions are mostly very short, between two days and a half and four days on average. This may explain the large share (6.7% to 35.2%) of workers aged 45–54 receiving training. Furthermore, around half of spells last for 1 or 2 days. Specifically, we see that between 1998 and 1999, the proportion of x-days training sessions for the 45–49 years old remained quite stable (whatever the number of days x is) in small firms, while in large firms the proportion of short spells strongly increased and the share of long spells (more than 4 days) decreased. However, these statistics are very difficult to interpret, first because the picture is not so clear-cut regarding workers aged 50–54 and also because

[19] To provide a more accurate picture of the effect of the Delalande tax reform on training decisions, we also estimate the effect for alternative types of training spells (see Section 3).

[20] Data have been merged at the individual level, using the individual identification number.

[21] The LFS also includes some information on training, but it is not detailed and precise enough. For instance, it would be impossible to know who pays for the training session, which is very important in our study.

[22] Another reason for excluding workers aged 55 and over is their much higher separation rate, compared to workers below 55. Using again the French LFS, we see that the separation rate among workers aged 54 over the period 1998–1999 equals 5.83%, while the separation rate among workers aged 55 over the same period equals 14.51%, that is almost three times higher. Because French older workers aged 55 and over were allowed in 1999 to benefit from unemployment insurance without job-search requirements until they reach the legal retirement age, employers used to consider unemployment benefits as early retirement schemes. This certainly explains why we observe higher separation rates once workers reached the age of 55.

[23] Following Hairault et al. (2010), the distance to retirement is captured by the difference between the current age and the retirement age. Considering the French pension system, the retirement age can be approximated by the required number of contributive years from graduation to get the full pension rate. The distance to retirement for an individual is then equal to the full pension age minus his or her current age. However, if a person enters the job market at a very young age, (s)he cannot retire before the eligibility age for the full pension (60 years old) even though (s)he has accumulated the required number of contributive quarters before this age. In this case, the retirement age is then set at 60 and the distance to retirement is 60 minus the current age. Furthermore, we account for the fact that at 65 individuals can draw a full pension whatever their number of contributive years.

[24] Even though it is possible to consider training spells occurring between January 1999 and March 2000 with the training survey, we limit the analysis to training sessions that took place between March and December 1999, in order to homogenize pre- and post-reform periods.

Table 2
Description of the sample, before and after the reform (in shares).

	Pre-reform		Post-reform	
	Large firms	Small firms	Large firms	Small firms
Training rate among the 45–49 considering:				
Any purpose (i to vi)[a], any sponsor	0.285	0.155	0.352	0.177
Any purpose (i to vi)[a], firm-sponsored	0.255	0.122	0.332	0.127
Purpose (i)[a], any sponsor	0.235	0.138	0.310	0.155
Purpose (i)[a], firm-sponsored	0.222	0.116	0.307	0.116
Training rate among the 50–54 considering:				
Any purpose (i to vi)[a], any sponsor	0.234	0.119	0.284	0.148
Any purpose (i to vi)[a], firm-sponsored	0.212	0.081	0.252	0.104
Purpose (i)[a], any sponsor	0.205	0.096	0.241	0.104
Purpose (i)[a], firm-sponsored	0.194	0.067	0.227	0.081
Average duration of training sessions (in days)				
Among the 45–49	3.43	2.95	2.77	2.62
Among the 50–54	3.5	3.78	2.81	3.64
Proportion of 1-day training sessions				
Among the 45–49	0.2	0.238	0.324	0.286
Among the 50–54	0.204	0.222	0.286	0.091
Proportion of 2-days training sessions				
Among the 45–49	0.2	0.238	0.225	0.238
Among the 50–54	0.204	0	0.222	0.273
Proportion of 3-days training sessions				
Among the 45–49	0.138	0.238	0.144	0.286
Among the 50–54	0.111	0.222	0.175	0.182
Proportion of 4-days training sessions				
Among the 45–49	0.088	0	0.72	0.048
Among the 50–54	0.056	0.111	0.111	0
Proportion of 5-days training sessions				
Among the 45–49	0.188	0.19	0.126	0.048
Among the 50–54	0.222	0.222	0.127	0.273
Proportion of 6-days training sessions				
Among the 45–49	0.188	0.095	0.108	0.095
Among the 50–54	0.204	0.222	0.079	0.182
Age				
45–49	0.565	0.573	0.455	0.472
50–54	0.435	0.427	0.545	0.528
Marital status				
In couple	0.894	0.87	0.906	0.861
Living alone	0.106	0.12	0.094	0.139
Nationality				
French	0.969	0.956	0.969	0.956
Others	0.031	0.044	0.031	0.044
Education				
No diploma	0.368	0.383	0.368	0.383
CAP-BEP	0.405	0.443	0.405	0.443
Baccalaureate	0.092	0.082	0.092	0.083
College degree	0.135	0.092	0.135	0.092
Distance to retirement				
<8 years	0.216	0.215	0.316	0.313
9–11 years	0.278	0.288	0.296	0.323
12–14 years	0.346	0.367	0.357	0.332
More than 14 years	0.16	0.13	0.031	0.032
Job seniority				
Less than 5 years	0.114	0.388	0.100	0.360
6–10 years	0.111	0.172	0.114	0.191
11–20 years	0.202	0.186	0.189	0.197
More than 20 years	0.573	0.255	0.595	0.249
Training plan in the firm				
Yes	0.736	0.271	0.736	0.271
No	0.264	0.729	0.264	0.729
Occupation				
Executives	0.175	0.133	0.180	0.139
Middle management	0.274	0.258	0.279	0.244
Employees	0.078	0.064	0.078	0.051
Blue collar workers	0.473	0.546	0.463	0.566
Sector				
Industry	0.590	0.266	0.595	0.272
Building	0.059	0.196	0.061	0.199
Services	0.351	0.538	0.344	0.529
Type of job				
Full-time	0.978	0.975	0.978	0.975
Part-time	0.022	0.025	0.022	0.025
Monthly net wage	1786.9	1527.2	1807.9	1529.8
Number of observations	639	316	639	316

1. Source: French Labour Force Survey (1998 and 1999) and training survey "Formation Continue 2000". Only male workers employed in the private sector are considered.
[a] The purpose of the training activity has to fit in one of the following categories: i) to adapt to the job, ii) to switch to another job or to get a job, iii) to obtain a diploma or a certification, iv) to execute political duties, v) personal or cultural reasons, and vi) no specific reason.

the number of observations is very small, particularly for small firms. Consequently, we will not estimate the effect of the 1999 reform on the length of training sessions and we will focus on the incidence of training.

Regarding the characteristics of our sample, comparing individuals employed in large vs. small firms is of particular interest to us. Not surprisingly, the probability that the firm has a training plan is much higher in large firms (about 74% in large firms against 26% in small establishments).[25] Similarly, we observe that job seniority tends to be much higher for individuals working in large firms than for those working in small firms. Furthermore, the sectoral composition strongly depends on firms' size as well. For instance, the building sector represents barely 6% of all the jobs in large firms while it is about 20% in small firms. Small firms are also characterized by the predominance of the tertiary industry while the manufacturing sector is the greatest in large organizations. Lastly, regarding the wage distribution, workers in large firms have on average better paid jobs than workers employed in small firms (about 1800 euros a month vs. 1528 euros).

5. Empirical analysis

5.1. Identification strategy

The goal of the paper is to measure the impact of stricter employment protection for older workers on firms' incentives to provide training. In order to do so, we exploit a discontinuity in the Delalande tax reform: in 1999 the tax for firms with at least 50 workers was increased while it remained unchanged for firms with less than 50 employees. The treatment was an unexpected one-off change in government policy[26] and applied almost equally to all members of the treatment group.[27] Furthermore, the one-off nature of the change makes it easy to select specific pre- and post-treatment points in time and apply difference-in-differences estimators (DiD hereafter). However, due to the differences in characteristics between treated and non treated workers, our strategy is to combine DiD with propensity score matching.[28]

Basically, we are going to compare the change in the average training propensity of workers employed in firms with fewer and with more than 50 employees (i.e. control vs. eligible groups), before and after the 1999 reform (i.e. pre- and post-treatment). More formally, let $P_{i,t}$ be a dummy variable equal to 1 if a worker i takes part in a firm-financed training session at time t, with $t \in \{1998; 1999\}$. Treatment and control groups are identified by the dummy variable T_i, such that $T_i = 1$ if the worker i is employed in a large firm (i.e. a firm with more than 50 employees). A set of covariates $X_{i,t}$ assumed to significantly affect the access rate to firm-sponsored training is also included. Specifically, we include common training determinants such as nationality, marital status, education, occupation, job seniority, part-time or full-time job, wage quartiles, sector of activity, existence of a training plan in the firm, as well as the distance to retirement. Because all these covariates are dummy variables, we then estimate the following linear probability model for each group/year:

$$E\left(P_{i,t} = 1|X_{i,t}, T_i, \tau_t\right) = \beta X_{i,t} + \delta_1 T_i + \delta_2 \tau_t + \alpha_{DiD}(T_i * \tau_t) + c_i + U_{i,t} \quad (9)$$

where c_i is an individual effect and τ_t a dummy variable that takes the value of 1 in 1999 (i.e. after the reform) and zero otherwise. The interaction term between the large firm dummy T_i and the post-reform dummy τ_t captures the effect of interest. Specifically, the coefficient α_{DiD} gives the average effect of treatment on the treated (ATT):

$$\begin{aligned}\alpha_{DiD} = &\left[E\left(P_{i,99} = 1|T = 1\right) - E\left(P_{i,98} = 1|T = 1\right)\right] \\ &- \left[E\left(P_{i,99} = 1|T = 0\right) - E\left(P_{i,98} = 1|T = 0\right)\right].\end{aligned} \quad (10)$$

It is worth noting that the DiD estimator allows us to remove unobservable individual-specific effects constant over time and common time effects.[29] Thus, the DiD estimator is based on the identifying assumption that, in the absence of the treatment, the treated and the non-treated do not have different time trends relative to their outcomes: this is the "time-invariance" assumption, which here formally writes:

$$\begin{aligned}&E\left(P_{i,99} = 1|T = 1, X_{i,99}\right) - E\left(P_{i,98} = 1|T = 1, X_{i,98}\right) \\ &= E\left(P_{i,99} = 1|T = 0, X_{i,99}\right) - E\left(P_{i,98} = 1|T = 0, X_{i,98}\right).\end{aligned} \quad (11)$$

In our setting, the time-invariance assumption means that the average training propensity of workers employed in large firms would have shifted after the reform identically to the training propensity of workers employed in small ones, had they worked in small firms too. In particular, this implies that being employed in a large firm should be similar to working in a small organization. However, as highlighted in Table 5 in Appendix A, one may not be confident with the time invariance assumption since job seniority, wage quartiles, sector and whether the firm has a training plan are significantly different in small and large organizations, both for the 45–49 and the 50–54 age groups (Columns 1 and 3). Because these variables may greatly matter in explaining access to firm-financed training, the observed deviations make both groups hard to compare. Therefore, we should estimate the causal effect of the change in the Delalande tax on workers' training propensity by accounting for these inherent differences in the distribution of covariates between the treated and the control groups. Accordingly, we choose to use a conditional difference-in-differences approach (CDiD hereafter), as first suggested by Heckman et al. (1997) and Heckman et al. (1998).[30] It allows us to remove systematic observed differences between workers employed in a small and those working in a large firm.

To implement the CDiD estimator, the first step requires us to build a suitable small firm control group through a propensity score matching approach, i.e. by matching each worker employed in a large firm with a similar worker in a small firm. Following Rosenbaum and Rubin (1983), the matching procedure is based on the propensity score, i.e. the conditional probability of working in a large firm given pre-treatment characteristics X: $e(X) = P(T_i = 1|X_i)$[31]. Once this conditional probability is estimated, it is then possible to construct an estimate of the average training propensity of workers employed in large firms, had they worked in a small firm (counterfactual), by taking a weighted

[25] This result is in line with Bassanini et al. (2007) and Montizaan et al. (2010) for instance, who show that training incidence is much higher within larger organizations.

[26] The tax reform was fixed by decree (No. 98-1201) on 28 December 1998, that is exactly three days before being applied (1 January 1999). So almost no firms employing more than 50 workers in 1998 were able to foresee the reform and reduce their size for 1 January 1999 on purpose.

[27] As shown in Table 1, even though the 1999 reform led to doubling the tax for most ages, it did not apply to all workers strictly equally (the rise was less significant for the oldest ages and the tax trebled in some cases). However, given the sizeable tax reform, the effect is never insignificant so that evaluating the impact on the entire group is not a problem.

[28] We should ideally exploit the exogenous threshold determining assignment to treatment — the 50 workers employed threshold — using the Regression Discontinuity method (see overviews papers from van der Klaauw (2008) and Imbens and Lemieux (2008) for instance) but we are unable to do so unfortunately. In our dataset, we only have information on establishment size through a categorical variable. More precisely, there are 9 categories: '1 or 2 workers in the firm', '3 to 5 workers', '6 to 9', '10 to 19', '20 to 49', '50 to 99', '100 to 499', '500 to 999' and 'more than 1000 workers in the firm'. Such a variable does not match the Regression Discontinuity Design de facto.

[29] While DiDs solve the problem of time-invariant unobservable factors, time variant unobserved heterogeneity possibly remains unidentified.

[30] Using data from the National Supported Work (NSW), Smith and Todd (2005) show that "the difference-in-differences matching estimator performs the best" among nonexperimental matched based estimators. They "attribute its performance to the fact that it eliminates potential sources of temporally invariant bias present in the NSW data, such as geographic mismatch between participants and nonparticipants…".

[31] Matching attempts to mimic randomization where covariates of treatment and control groups are balanced on average. For a complete discussion of matching methods, see Dehejia and Wahba (2002) and Caliendo and Kopeinig (2008).

average of the training propensity of workers in small firms. More specifically, we apply kernel matching estimators to match observations.[32] Unbiased estimates can be obtained only if the conditional independence assumption is satisfied, which implies that, conditionally on some set of observed characteristics, training propensities of workers employed in small firms are independent of the fact of being employed in a large firm:

$$E(P_{i99}=1|T_i=0)-E(P_{i98}=1|T_i=0)\perp T_i|e(X_i). \quad (12)$$

In a second stage, we estimate ATT using a DiD regression and weighting non-treated observations according to their propensity score. To take into account that weights are estimated as well, the standard errors of the impact estimates are calculated by bootstrap using 1000 replications for each estimate.[33] Denoted α_{CDiD}, the DiD matching estimator is written as:

$$\begin{aligned}\alpha_{CDiD} = \Sigma_i[&\left(E\left(P_{i,99}=1|T_i=1\right)-E\left(P_{i,98}=1|T_i=1\right)\right)\\ &-\Sigma_j W_{ij}\left(E\left(P_{j,99}=1|T_i=0\right)-E\left(P_{j,98}=1|T_i=0\right)\right)]\end{aligned} \quad (13)$$

where W_{ij} is the weight placed on comparison observation j for individual i.

5.2. Results

5.2.1. Matching

In our application of the CDiD estimator, we first use a probit model to estimate the probability of working in a large firm, for both age groups (45–49 and 50–54). We can be confident in the matching procedure since the fit of the probits is good: on average, they correctly predict the treatment status in more than 75% of cases.[34] The results of the probit estimates are shown in Appendix B (Table 6). Not surprisingly, some observables such as job seniority or the presence of a training plan in the firm strongly affect the probability of working in a firm with 50 workers or more. Workers employed in the industrial sector are also more likely to work in a large establishment.

In addition, propensity score matching can be successful only if the estimated propensity scores of workers employed in small and large firms overlap sufficiently. The most straightforward way to check the common support between treatment and control groups is a visual analysis of the density distribution of the propensity score in both groups. Fig. 1 shows kernel density estimates of the propensity scores for workers employed in a small firm and for those employed in large firms, by age group. First, the left-hand side graphs (i.e. before matching) show that for both age groups, the overlapped region is largely wide: almost all treated workers have at least one non-treated worker with the same value of the propensity score. Accordingly, adopting the Min–Max method suggested by Dehejia and Wahba (2002) does not imply eliminating a relevant number of observations.[35] Secondly, we observe that for high values of the propensity score, the proportion of non-treated workers to be matched with treated workers becomes quite small. To control for the fact that in that case, matching is performed on too small a number of non-treated observations, we combine the Min–Max method with trimming (Smith and Todd (2005)).[36]

Furthermore, since we do not condition on all covariates but on the propensity score, it has to be checked whether the matching procedure is able to balance the distribution of the relevant variables in both groups (treated and controls). Table 5 in Appendix A shows differences in means of covariates before and after conditioning. Differences in the mean for job seniority, whether the firm has a training plan and wages quartiles are no longer significant after matching. Significant differences for the sector remain but the matching procedure led to strongly decreasing the differences in means. Comparing left- and right-hand side graphs in Fig. 1 confirms that the matching procedure made comparable the distribution of covariates across groups.

6. Findings

Table 3 shows estimates of the average effect of the increase in the Delalande tax scheme in 1999 on the incidence of training among workers employed in a large firm (ATT), by age group. The first line corresponds to the estimates from a simple DiD approach (without matching) while the other estimates below are obtained combining DiD with different matching techniques. As a baseline, we use a Gaussian Kernel Estimator setting the bandwidth to 0.06. As the use of kernel matching estimators may increase the bias, especially when the bandwidth is high, we report kernel matching estimates with a lower bandwidth, equal to 0.005. We also perform nearest neighbor propensity score matching with replacement and one single neighbor, setting a caliper equal to 0.01 so that the propensity score value of the nearest neighbor is close to the value of the treated individual. In addition, we test the matching specification by using a more restrictive definition of the common support. Specifically, we combine a Min–Max approach with trimming and discard treated units with a propensity score corresponding for the non-treated workers to the lowest 10% of the density distribution. As shown in Table 3, the ATT for the 45–49 age group equals 0.092 with a simple DiD approach and ranges from 0.097 to 0.129 for the different CDiD estimates. In addition, all coefficients are at least significant at the 5% level. Conversely, the DiD coefficient for the 50–54 age group equals 0.017 and CDiD estimates range from −0.007 to 0.035. None is statistically significant at the 10% level (whatever the definition of the variance). Finally, we show the estimates of the ATT using exact matching methods on some covariates.[37] For the 45–49 age group, 62.32% of treated individuals are matched (we lose one third of the treated group). The ATT is 8.5 percentage points and is significant at the 5% level but only if we consider the asymptotic conditional variance *à la*

[32] Gaussian Kernel Matching is a non-parametric matching estimator that uses weighted averages of all workers in the control group to generate the counterfactual outcome. Weights depend on the distance between each worker from the control group and the treated observation for which the counterfactual is estimated. More formally, let p_i be the propensity score of the treated individual i and p_j the propensity score for the non-treated individual j, the weight placed on observation j is defined by $K\left(\frac{p_i-p_j}{h}\right)$, where K is the kernel estimator (here we consider the normal density function) and h represents the bandwidth. The higher the bandwidth, the lower the variance but the higher the bias. One major advantage of the Gaussian Kernel Matching approach is the smaller variance, which is achieved because more information is used. However, a drawback is that observations which are bad matches may also be used. As Caliendo and Kopeinig (2008) show, this could increase the bias (which can be reduced by reducing the bandwidth). For the sake of robustness, we will also show results obtained with other matching techniques, e.g. nearest neighbor matching with replacement and exact matching on some covariates.

[33] Bootstrapping techniques to estimate standard errors of matching estimators are widely applied (Caliendo and Kopeinig (2008)). For the exact matching method, we use the approach of Abadie and Imbens (2006) who show that the standard bootstrap is not valid in this case.

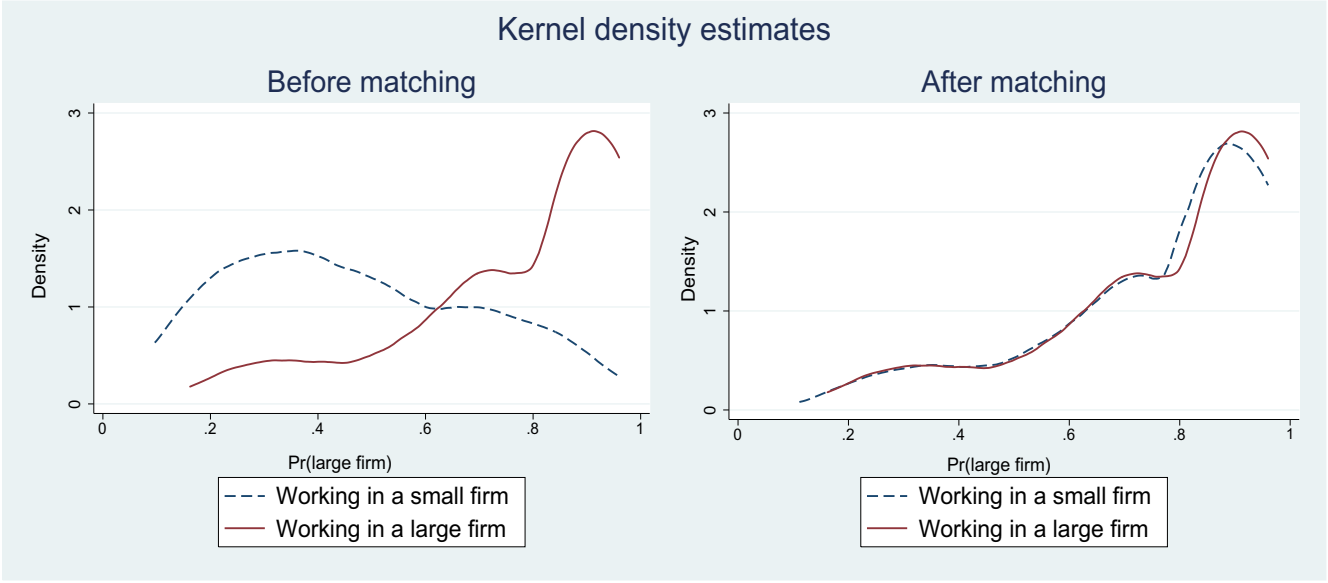
[34] More precisely, probits regression of the propensity score predicts the treatment status for 78.41% of the 45–49 age group and for 75.37% of the 50–54 age group.

[35] More precisely, implementing the common support condition leads us to discard 18 observations of the 45–49 age group and 27 of the 50–54 age group. See Table 3.

[36] The results obtained with this new definition of the common support are shown in Table 3.

[37] We consider the covariates that differ strongly across groups, i.e. having a training plan in the firm, job seniority, wage, occupation and sector. Furthermore, the drawback of this approach is referred to as a dimensionality curse: the higher the number of variables we consider for matching, the lower the number of treated units that can be matched. However, it ensures the removal of the whole selection bias on these observables. In addition, as Abadie and Imbens (2006) argue, bootstrapping is not valid to estimate standard errors with the exact matching method. Therefore, we follow their approach to compute the asymptotic variance. The idea is first to match treated units to treated units and control units to control units. Then for each observation i, we estimate the conditional variance for a fixed number of neighbors (here we consider one single neighbor) $\hat{\sigma}^2(X_i,W_i)$ as: $\hat{\sigma}^2(X_i,W_i)=\frac{1}{2}(Y_i-Y_j)^2$, where j is the closest individual to the individual i. The conditional variance of the ATT is then given by $\hat{V}=\frac{1}{N_1}\sum_{i=1}N(T_i-(1-T_i)K_M i)^2\hat{\sigma}^2(X_i,W_i)$, where N_1 corresponds to the number of treated units and $K_M i$ stands for the number of times unit i is used as a match given that M matches per unit are used.

(a) For the 45-49 age group



(b) For the 50-54 age group

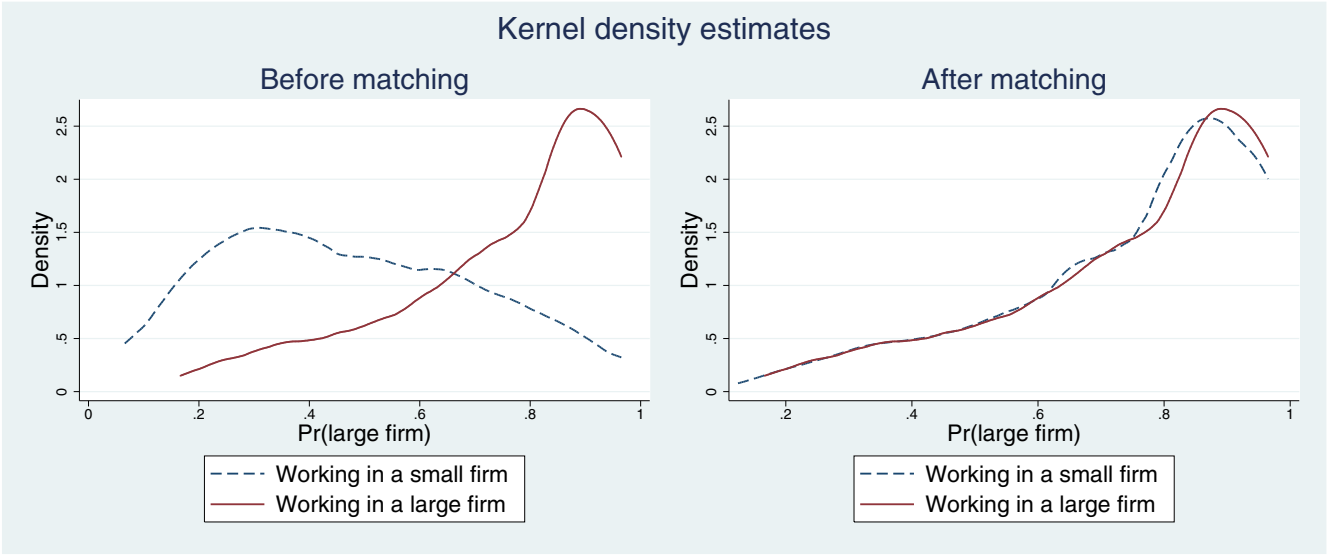


Fig. 1. Common support of the propensity scores.

Abadie and Imbens (2006). Moreover, the loss of information is much higher for workers aged 50–54 years (only 30.93% of matched treated units). The ATT is negative and not statistically significant, but these estimates should be interpreted with caution given the small sub-sample of treated units that we consider.

Specifically, our findings show that the rise in the tax in 1999 led to a significant and substantial increase in the training propensity of the 45–49 age group employed in a large firm. As illustrated by our intuitive theoretical predictions, this strong and positive effect shows how complementary the job destruction and training investment decisions of firms actually are. The idea is that the increased firing tax after 50 changes the firing policy of firms, both below 50 and after 50. Hence, knowing that dismissals will cost much more after the age of 50, firms can decide to engage in training activities before the worker turns 50 to increase productivity and job value, which in turn will lower the probability of a job destruction after 50 (which would be costly because of the tax increase).

Secondly, we find a small and not significant average effect of the tax reform on the incidence of training for the 50–54 age group. Even though the 1999 reform sizeably decreased layoffs of the 50–54 age group in large firms (Behaghel et al., 2008), this did not lead to a significant increase in their training propensity. In line with our theoretical illustration and existing literature, this result reflects the positive correlation between the horizon of a worker and her or his probability of benefiting from a firm-provided training session.[38]

[38] One may rightly question whether workers aged 50–54 are really close to the retirement age. Indeed, considering the legal retirement age at 60 in France, workers aged 50–54 are not that close to retirement at first sight. However, because unemployed older workers aged 55 and over were entitled in 1999 to unemployment benefits without job-search requirements until they reach the legal retirement age, employers tended to use the unemployment insurance as pre-retirement schemes (Gruber and Wise, 1998). Consequently, we can with good reason consider that workers aged 50–54 are close to the exit age. Again, the separation rate among workers aged 55 over the period 1998–1999 was almost three times higher than the separation rate of workers aged 54 (see footnote number 22).

Table 3
Effect of the tax increase in 1999 on the incidence of training among workers employed in large firms (ATT).

	45–49 year-old		50–54 year-old	
DiD[a] (without matching)	0.092** (0.043)		0.017 (0.045)	
Matching procedure	CDiD[b]	S.E.[c] (bootstrapped)	CDiD[b]	S.E.[c] (bootstrapped)
Common support (Min–Max)	18 treated units off support		27 treated units off support	
Kernel (bandwidth = 0.06)	0.097**	0.037	0.025	0.053
Kernel (bandwidth = 0.005)	0.104**	0.052	0.003	0.072
Nearest neighbor (caliper = 0.01)	0.122*** (93% of matched treated)	0.048	0.035 (95.8% of matched treated)	0.066
Common support (trimming: 10%)	34 treated units off support		37 treated units off support	
Kernel (bandwidth = 0.06)	0.101***	0.036	0.008	0.05
Kernel (bandwidth = 0.005)	0.111**	0.052	−0.007	0.053
Nearest neighbor (caliper = 0.01)	0.129*** (92.67% of matched treated)	0.048	0.017 (100% of matched treated)	0.064
Exact matching (13 variables)	0.084** (62.32% of matched treated)	0.039[d]	−0.035 (37.93% of matched treated)	0.046[d]
Number of observations	542		413	
Among the high-skilled[e]	0.068	0.067	0.012	0.119
Number of observations	224		186	
Among the low-skilled[e]	0.099*	0.059	0.059	0.065
Number of observations	318		227	

Significance level: ***p < 0.01, **p < 0.05, *p < 0.1.

[a] DiD coefficients were obtained regressing the dependent variable on the same set of regressors as for the estimation of propensity score and also on a dummy equal to one if the individual works in a large firm, a dummy equal to one if the dependent variable is observed during the post-reform period and an interaction term between these last two dummies. The ATT effect is given by the coefficient associated with the interaction term. Standard errors are in parentheses.

[b] Dummies for nationality, marital status, distance to retirement, diploma, job seniority, occupation, part-time jobs, sector, wages quartiles and a dummy for the presence of a training plan in the firm are included in the set of regressors.

[c] Standard errors calculated by bootstrap using 1000 replications for each estimate.

[d] Sample standard errors (Abadie and Imbens, 2006).

[e] CDiD coefficients were obtained using kernel matching estimators with a bandwidth equal to 0.06 and a common support defined by the Min–Max approach. To be precise, 9 (31) treated units are off support for the high-skilled workers aged 45–49 (50–54), 14 (34) for the low-skilled workers aged 45–49 (50–54).

Moreover, as argued in the theoretical illustration, the new Delalande tax schedule in 1999 may have had heterogeneous effects among the 45–49 age group according to the skill level, depending on how fast technology changes. Specifically, when the productivity level is known for sure (persistent idiosyncratic productivity), the lower the worker's productivity, the higher the risk of dismissal, hence the greater the training incentives (to increase the job value after 50). Accordingly, we should expect a greater effect of the tax reform on the incidence of training for the low-skilled workers than for the high-skilled. On the other hand, in the event of a rapid change of technology, there is uncertainty about the productivity level and the job value in the future (after 50). Hence, firms have incentives to train the 45–49 year olds, whatever the worker's specific human capital level. Therefore, we should not find any significant differences across skill groups.

We address this concern by showing in Table 3 results based on the estimation of Eqs. (9) and (13) and classifying by occupation. More precisely, we consider two groups of workers: high-skilled and low-skilled workers.[39] Estimates do show heterogenous effects according to the skill level. In particular, the incidence of training among the low-skilled significantly rose by about 10 percentage points after the tax reform, while the increase was smaller (+6.8 percentage points) and not statistically significant for the high-skilled.

6.1. Robustness and sensitivity

Table 3 showed that the incidence of training among the 45–49 age group employed in large firms has increased. Other reforms than the new Delalande tax schedule (and promulgated during the same time period as well) could also have contributed to this rise. In particular, this would be the case if some institutional changes affected small and large firms differently, because the CDiD approach only solves the problem of common time effects assuming that changes are identical between the treatment and control groups. Among some reforms likely to affect firms' incentives to sponsor training, the introduction of the 35-hour working week regulation in 1999 may be a good candidate. As shown by Aeberhardt et al. (2012), the signing date of the agreement on working time reduction was strongly dependent on firms' size, since larger establishments reduced their working time first. Consequently, this regulation could have had a differentiated impact on firms' training incentives between the treated and the controls. As we have information on the effective working hours of respondents, we suggest only considering individuals who did not shift to the 35-hour working week (i.e. those who reported a number of weekly working hours of 39 or more). The results are presented in Table 4. First, a small proportion of individuals experienced a reduction in their working time (around 16%). Secondly, we find a significant ATT of 9.3 percentage points (CDiD estimate), which is very close to the baseline estimate (9.7 percentage points). So, it seems that our results remain unchanged if we put aside the 35-hour working week change.

In addition, we suggest another simple test to check whether any other change occurred during the time period 1998–1999 which could have contributed to the effects we highlight. Specifically, we suggest looking at the ATT for younger cohorts. Under the null hypothesis, the effect of another regulation would be the same between groups and would be removed by the DiD. Therefore, the estimates should be significant only for the 45–49 age group. In the alternative hypothesis, the effect of another regulation would differ across groups and the estimates should be significant for all cohorts of workers. To perform this

[39] The group of high-skilled workers is made up of executives ("cadres" in French) and middle management ("professions intermédiaires"), while low-skilled workers include employees ("employés") and blue collar workers ("ouvriers").

Table 4
Robustness of the ATT (for the 45–49).

ATT	CDiD[a]	S.E. (bootstrapped)	DiD[b]	S.E.
Among workers at 39 hours per week	0.093**	(0.040)	0.086**	(0.041)
Number of observations	454			
Number of treated units off support	13			
Among the 25–29	−0.048	(0.065)	0.012	(0.055)
Number of observations	320			
Number of treated units off support	28			
Among the 30–44	−0.042	(0.038)	−0.016	(0.023)
Number of observations	1697			
Number of treated units off support	10			
Changing the treatment variable	−0.026	(0.100)	−0.078	(0.092)
Number of observations	319			
Number of treated units off support	14			
Purpose (i)[c], firm-sponsored (baseline)	0.097**	(0.037)	0.086**	(0.04)
Any purpose (i to vi)[c], firm-sponsored	0.08*	(0.041)	0.07*	(0.041)
Purpose (i)[c], any sponsor	0.077*	(0.041)	0.058	(0.041)
Any purpose (i to vi)[c], any sponsor	0.006	(0.056)	0.041	(0.043)
Number of observations	542			
Number of treated units off support	18			

Significance level: **p < 0.05, *p < 0.1.

[a] The CDiD coefficients have been obtained using kernel matching estimators with a bandwidth equal to 0.06 and a common support defined by the Min–Max approach. Dummies for nationality, marital status, distance to retirement, diploma, job seniority, occupation, part-time jobs, sector, wages quartiles and a dummy for the presence of a training plan in the firm are included in the set of regressors. Standard errors are calculated by bootstrap using 1000 replications for each estimate.

[b] The DiD coefficients have been obtained regressing the dependent variable on the same set of regressors as for the estimation of propensity score and also on a dummy equal to one if the individual works in a large firm, a dummy equal to one if the dependent variable is observed during the post-reform period and an interaction term between these last two dummies. The ATT effect is given by the coefficient associated with the interaction term.

[c] The training survey describes the purpose of the training activity, which has to fit into one of the following categories: i) to adapt to the job, ii) to switch to another job or to get a job, iii) to obtain a diploma or a certification, iv) to execute political duties, v) personal or cultural reasons, and vi) no specific reason.

test, we run new DiD and CDiD estimations based on Eqs. (9) and (13) but without any selection on age, including workers aged between 25 and 44. The results are presented in Table 4. They show that the null hypothesis cannot be rejected. Specifically, estimates for workers aged 25–29 and for workers aged 30–44 are small in magnitude and not statistically significant.

Furthermore, one may argue that the rise in the incidence of training for the 45–49 age group employed in large firms could simply come from a heterogeneous effect between small and large firms, instead of strictly being the result of the change in the tax schedule in 1999. If that is true, besides disparities between small and large firms, there should also be a difference in training propensities between large firms and very large firms. But if it is not true, then the threshold of 50 employees in the firm should be the key determinant. Therefore, we suggest another test to check whether the new tax schedule in 1999 was really responsible for the significant difference in the training propensity we have found for the 45–49 age group. To be precise, we perform new DiD and CDiD matching estimations based on Eqs. (9) and (13), but changing the treatment variables. We now only include workers employed in a firm with more than 50 employees and consider that the treatment group is composed of workers employed in firms with more than 100 workers while the control group is made up of workers in firms employing between 50 and 99 workers. The idea is to have only treated observations in reality in the sample (i.e. workers employed in a firm with at least 50 workers). As can be seen in Table 4, there are no longer any significant training disparities between firms. It just goes to show that the threshold of 50 employees in the firm matters above all, and not the establishment size per se. Therefore, we can be confident enough in our ability to recover the effects of the new Delalande tax schedule in 1999.

Lastly, the DiD and CDiD estimates presented in Table 3 only concern the firm-financed training spells aiming at adapting to the job. However, we can also look at the incidence of training in pre- and post-reform periods among the 45–49 age group, for alternative types of training. Specifically, we perform new DiD and CDiD matching estimations based on Eqs. (9) and (13) but considering any sponsor and any purpose of the training sessions. As shown in the last part of Table 4, the ATT is greatest when considering firm-financed training spells aimed at adapting to the job (our baseline definition). Moreover, it seems that the characteristic of 'firm-sponsored' is more responsible for the strong and positive ATT than the characteristic of 'aimed at adapting to the job' (the estimates are a bit higher for firm-financed training sessions without specifying any aim than for sessions aimed at adapting to the job without specifying any sponsor, and the estimate from the simple DiD is not significant in the third specification). Conversely, the ATT is almost nothing when neither purpose nor sponsor is specified (but such training spells are not very common in our sample).

7. Conclusion

This paper analyzes the effect of stricter employment protection among older workers on firms' incentives to finance training. We first present theoretical illustrations of the possible effects and then use French data in a quasi-natural experiment framework to give empirical results. In particular, we exploit the unexpected increase in the Delalande firing tax schedule in 1999 that concerned only firms employing at least 50 workers. We implement a conditional difference-in-differences estimator to remove selection bias on treatment of observables, individual specific effects constant over time and macro-effects common to both groups.

We first find that the new Delalande tax schedule for large firms significantly raised the incidence of employer-provided training for workers aged 45 to 49, by about 10 percentage points. As our simple theoretical model shows, this result comes from a complementarity effect between firing and training decisions of firms. The idea is that the increased firing tax after age 50 changes the firing policy of firms, both for workers under and over 50. Therefore, knowing that dismissals will cost much more after age 50, firms have strong incentives to engage in training activities before the worker turns 50 to increase productivity and job value, which in turn will lower the probability of a job destruction after 50 (which would cost much more because of the tax increase). However, for workers over 50 employed in large firms, the new tax schedule had no significant impact. This is due to the proximity of retirement. Furthermore, a skill-breakdown of the effect among the 45–49 age group shows that the new tax schedule had a bigger effect on the incidence of training among low-skilled workers. Low-skilled jobs require more training (to increase the job value after 50) since the lower the worker's productivity, the higher the risk of dismissal.

While tackling an interesting research issue, this paper has several limits. First and foremost, we only highlight a short-term effect of the rise in the Delalande tax schedule. Indeed, the data only allow for a one year follow-up after the reform. Hence, the new tax schedule may have had a considerable impact in the first year but a much weaker effect afterwards. To be able to give long-term predictions, we should rather rely on a structural estimation of a model with a finite life time horizon à la Chéron et al. (2011), extended to account for training investments by firms. This would allow us to better understand how training and firing decisions of firms about older workers are related. We should also be able to analyze the effect of implementing training subsidies as a way of increasing older workers' access to training.[40] These concerns are on our research agenda.

[40] As we mentioned in the Introduction Section, Picchio and van Ours (2011) show that better access to on-the-job training has an effect on the employability of workers, even for older workers.

Appendix A. Differences in means of covariates between control and treated groups

Table 5
Differences in means of covariates between control and treated groups.

	45–49		50–54	
	Before matching	After matching	Before matching	After matching
Distance to retirement				
<8 years			0.007	0.051
9–11 years	−0.011	0.024	0.043	−0.031
12–15 years	0.059	0.037	−0.039**	−0.02
More than 14 years	−0.048	−0.06	−0.011	
Occupation				
Executives	−0.017	−0.042	−0.09**	0.018
Middle management	−0.006	0.053	−0.018	−0.036
Employees	−0.02	0.002	−0.016	0.052
Blue collar workers	−0.043	0.032	0.124***	−0.034
Job seniority				
Less than 5 years	0.245***	−0.019	0.113***	0.015
6–10 years	0.069**	−0.01	0.107***	0
11–20 years	−0.02	−0.034	0.016	0.028
More than 20 years	−0.294***	0.063	−0.263***	-0.043
Training plan	−0.46***	0.01	−0.46***	0.008
Sector				
Industry	−0.344***	−0.079**	−0.298***	−0.021
Building	0.119***	−0.005	0.161	0.003
Services	0.226***	0.084**	0.137***	0.018
Wages quartiles				
1st quartile	0.124***	0.029	0.195***	−0.015
2nd quartile	0.022	−0.011	−0.024	−0.011
3rd quartile	−0.064	−0.002	0.092***	0.004
4th quartile	−0.081*	−0.017	−0.155***	−0.022

Significance levels: *: 10%, **: 5%, ***: 1%.
Note: Only covariates for which the null hypothesis of equality of means can be rejected at a 10% level (at least) are presented.
Source: Labour Force Survey and Training Survey (1998 wave).

Appendix B. Propensity score estimation

Table 6
Propensity score estimation (Probit regressions).

Variable	45–49		50–54	
	Coeff.	Std. err.	Coeff.	Std. err.
In couple	−0.053	(0.213)	0.286	(0.226)
No French citizenship	0.004	(0.356)	−0.208	(0.373)
Education (ref: no diploma)				
CAP-BEP	−0.273*	(0.154)	−0.027	(0.168)
Baccalaureate	−0.191	(0.254)	−0.213	(0.297)
College degree	−0.03	(0.288)	−0.382	(0.346)
Distance to retirement				
<8 years			Ref	
9–11 years	Ref		0.075	(0.153)
12–14 years	−0.29	(0.191)	1.009*	(0.62)
More than 14 years	−0.098	(0.222)	.	.
Job seniority (ref: 5 years or less)				
6–10 years	0.165	(0.210)	0.154	(0.264)
11–20 years	0.425**	(0.192)	0.34	(0.261)
More than 20 years	0.58***	(0.173)	0.666***	(0.239)
Occupation (ref: workers)				
Executives	−0.184	(0.296)	−0.261	(0.331)
Middle management	−0.155	(0.18)	−0.277	(0.208)
Employees	0.538**	(0.256)	0.351	(0.335)
Existence of a training plan	1.041***	(0.137)	1.068***	(0.169)
Sector (ref: building)				
Industry	1.069***	(0.223)	1.075***	(0.242)
Services	0.298	(0.222)	0.441*	(0.246)
Part-time job	0.189	(0.429)	0.51	(0.498)
Wage quartiles (ref: 1st quartile)				
2nd quartile	0.081	(0.174)	0.251	(0.213)
3rd quartile	0.184	(0.196)	0.157	(0.235)
4th quartile	0.182	(0.257)	0.411	(0.302)
Intercept	−0.736**	(0.342)	−1.531***	(0.402)
Number of observations	542		410	
Pseudo R^2	0.259		0.234	

(continued on next page)

Source: Labour Force Survey (1998 wave) and Training Survey.
Significance levels: *: 10%, **: 5%, ***: 1%.

References

Abadie, A., Imbens, G.W., 2006. Large sample properties of matching estimators for average treatment effects. Econometrica 74 (1), 235–267.
Acemoglu, D., 1997. Training and innovation in an imperfect labour market. Rev. Econ. Stud. 64 (3), 445–464.
Acemoglu, D., Pischke, J.-S., 1998. Why do firms train? Theory and evidence. Q. J. Econ. 113, 79–119.
Aeberhardt, R., Givord, P., Marbot, C., 2012. Spillover Effect of the Minimum Wage in France: An Unconditional Quantile Regression Approach. Document de travail G2012-07 Institut National de la Statistique et des Études Économiques, Direction des Études et Synthèses Économiques.
Bassanini A., Booth A., Brunello G., De Paola M., and Leuven E. 2007. In Workplace training in Europe, edited by Wasmer E. (eds) Education in Brunello G., Garibaldi P. and Training in Europe. Oxford University Press.
Becker, G., 1962. Investment in human capital: a theoretical analysis. J. Polit. Econ. 70, 9–49.
Behaghel, L., 2007. La protection de l'emploi des travailleurs âgés en France: Une étude de la contribution Delalande. Ann. Econ. Stat. 85, 41–80.
Behaghel, L., Crépon, B., Sédillot, B., 2008. The perverse effects of partial employment protection reform: the case of French older workers. J. Public Econ. 92 (3–4), 696–721.
Belot, M., Boone, J., van Ours, J., 2007. Welfare-improving employment protection. Economica 295.
Cahuc, P., Ferracci, M., Zylberberg, A., 2011. Formation professionnelle: pour en finir avec les réformes inabouties. Technical Report. Institut Montaigne (Octobre).
Caliendo, M., Kopeinig, S., 2008. Some practical guidance for the implementation of propensity score matching. J. Econ. Surv. 22 (1), 31–72.
Chéron, A., Rouland, B., 2011. Inefficient job destructions and training with hold-up. Labour 25 (4), 397–420.
Chéron, A., Hairault, J.-O., Langot, F., 2011. Age-dependent employment protection. Econ. J. 121 (557), 1477–1504.
Dehejia, R., Wahba, S., 2002. Propensity score-matching methods for nonexperimental causal studies. Rev. Econ. Stat. 84 (1), 151–161.
Delort, A., 2013. La dépense nationale pour la formation professionnelle continue et l'apprentissage, en 2011. DARES Analyses (no. 080 (Décembre)).
Fella, G., 2005. Termination restrictions and investment in general training. Eur. Econ. Rev. 6 (49), 1479–1499.
Gruber, J., Wise, D., 1998. Social security and retirement: an international comparison. Am. Econ. Rev. 88 (2), 158–163.
Hairault, J.-O., Langot, F., Sopraseuth, T., 2010. Distance to retirement and older workers' employment: the case for delaying the retirement age. J. Eur. Econ. Assoc. 8 (5), 1034–1076.
Heckman, J., Ichimura, H., Todd, P., 1997. Matching as an econometric evaluation estimator: evidence from evaluating a job training programme. Rev. Econ. Stud. 64 (4) (605–542).
Heckman, J., Ichimura, H., Todd, P., 1998. Matching as an econometric evaluation estimator. Rev. Econ. Stud. 65 (2), 261–294.
Imbens, G., Lemieux, T., 2008. Regression discontinuity designs: a guide to practice. J. Econ. 142 (2), 615–635.
Khaskhoussi, T., Langot, F., 2008. La proximité de l'âge de la retraite. Quels effets sur l'accumulation du capital humain et sur la recherche d'emploi? Revue Économique 59 (3), 641–652.
Labruyère, C., 2005. Vocational certification: how the social partners contribute to developing the supply. Train. Employ. 65.
Lechthaler, W., 2009. The interaction of firing costs and firm training. Empirica 36 (3), 331–350.
Leuven, E., 2005. The economics of private sector training: a survey of the literature. J. Econ. Surv. 19 (1), 91–111.
Montizaan, R., Cörvers, F., De Grip, A., 2010. The effects of pension rights and retirement age on training participation: evidence from a natural experiment. Labour Econ. 17 (1), 240–247.
OECD, 2006. Live longer, work longer. Ageing and Employment Policies. OECD Publishing.
Picchio, M., van Ours, J.C., 2011. Retaining through training: even for older workers. IZA Discussion Paper No. 5591.
Pierre, G., Scarpetta, S., 2004. Employment regulations through the eyes of employers: do they matter and how do firms respond to them? IZA Discussion Paper No.1424.
Rosenbaum, P., Rubin, D., 1983. The central role of the propensity score in observational studies for causal effects. Biometrika 70, 41–55.
Schnalzenberger, M., Winter-Ebmer, R., 2009. Layoff tax and employment of the elderly. Labour Econ. 16 (6), 618–624.
Smith, J., Todd, P., 2005. Does matching overcome LaLonde's critique of nonexperimental estimators? J. Econ. 125 (1–2), 305–353.

van der Klaauw, W., 2008. Regression-discontinuity analysis: a survey of recent developments in economics. Labour 22 (2), 219–245.
Wasmer, E., 2006. General versus specific skills in labor markets with search frictions and firing costs. Am. Econ. Rev. 811–831 (no. 3).
Young, D., 2003. Employment protection legislation: its economic impact and the case for reform. Economic paper 186, European Commission Directorate-General for Economic and Financial Affairs.