Train More, Work Longer

The Effects of Pension Reforms and Training Incentives on the Retirement and Training Decisions of Older Workers

Erich Battistin (Padova, IRVAPP and IZA)
Giorgio Brunello (Padova, IZA, CESifo and ROA)*
Simona Comi (Milano Bicocca, IEIL and CHILD)
Daniela Sonedda (Piemonte Orientale and CRENoS)

This draft: 28 August 2012
PRELIMINARY AND INCOMPLETE

Abstract
We exploit the overlapping exogenous variation in both mandatory minimum early retirement age and government regional training subsidies in Italy during the 1990s and early 2000s to estimate the causal effect of training on the decision to retire and the causal impact of the distance from minimum retirement age – the horizon effect - on training. We find that both training and pension policies have contributed to reduce retirement and to increase training in our sample of Italian males aged 45 to 56. The size of the estimated effects is rather different, however. On the one hand, we find that an exogenous one - year increase in mandated minimum early retirement age has reduced the probability of retirement and increased training by close to 9 and 12 percent respectively. On the other hand, we show that one additional real euro per head spent in training incentives has reduced retirement and increased training by 0.3 and 2 percent. Ceteris paribus, it would take an additional 6 euro per head per year (close to 240 million euro) for training policies to be as effective as retirement policies in encouraging training among older workers. We provide evidence that higher training substantially reduces the probability of retirement in our sample, and that a longer working horizon significantly increases the individuals stock of training. We investigate reasons why training is so effective in reducing retirement. Our tentative estimates suggest that additional training reduces the risk of unemployment and increase earnings.

Acknowledgements: we are grateful to the participants at the Kansai Labor Seminar in Osaka for comments and suggestions. Part of this research was conducted while Giorgio Brunello visited ISER Osaka. The ILFI (Indagine longitudinale sulle famiglie italiane, 1997-2005) survey is the result of the joint efforts of Università di Milano Bicocca, Università degli Studi di Trento, Università degli Studi di Bologna (scientific coordinator: A. Schizzerotto). The usual disclaimer applies.

*Corresponding author: Department od Economics and Management, University of Padova, via del Santo 33, 35100 Padova, Italy; e-mail: giorgio.brunello@unipd.it
Introduction

Population ageing is a key challenge facing OECD economies. Over the next 50 years, all OECD countries will experience an important increase in the share of elderly persons in the population and a significant decline in the share of the population of prime working age (OECD, 2006). Policy options to offset ageing include the promotion of immigration, of higher fertility and faster productivity growth. Since many individuals aged 50 plus are out of the labour market, an additional policy option is to improve their employment prospect, so that they can stay longer in the labour market.

In 2010 close to 61% of the individuals aged 50 to 64 in OECD countries had a job, compared to 64% in North America and to 58% in Europe (source: OECD Labour Force Statistics). Policies that increase these activity rates can reduce the pressure of ageing on public finances, and at the same time protect living standards. The range of options includes reforms of retirement systems, which increase the option value of staying in the labour market rather than retiring, age dependent employment protection (see Cheron, Hairault and Langot, 2011), which increases the penalties faced by firms that layoff older workers, and training policies.

Many OECD Governments have embraced training as one possible means of bringing older workers out of unemployment or inactivity and into employment (see Mahyew and Rjkers, 2004). Training policies are often advocated because of the evidence showing that skilled older workers remain in the labour market longer than their unskilled peers. This evidence also suggests that there is a positive and statistically significant correlation across (European) countries between the incidence of training among older workers - relative to younger cohorts - and the average effective age of retirement (OECD, 2006, and Bassanini et al., 2007). Another important reason for encouraging training is that it might facilitate

1 The European Commission has strongly encouraged member countries to promote lifelong learning and training of older workers, by promoting equal opportunity in the workplace and by providing training incentives with the European Social Fund. See for instance Commission of the European Communities, 2002 and the European Directive on Equal Treatment of November 2000. According to the Bruges Communiqué, 2011, “...the future European labour market will be simultaneously confronted with an ageing population and shrinking cohorts of young people. As a result, adults - and in particular, older workers - will increasingly be called upon to update and broaden their skills and competences through continuing VET...” (p.2). US training policies targeted at older workers are reviewed by Eyster, Johnson and Toder (2008).
employability in the presence of technical shocks that depreciate existing skills (see Behaghel, Caroli and Roger, 2011).

This emphasis on training as a viable policy to address the problems of an ageing society has been met with some scepticism by economists. In particular, Heckman (2000) and Cunha et al. (2006) have argued that investing in the training of older workers is unlikely to yield high returns, as these workers and their employers have only a short time to recoup their investment and cannot benefit as much as younger workers from the dynamic complementarities that characterize human capital accumulation. An implication of this view is that, if one wants to promote the training of older workers, offering monetary incentives to workers and firms in order to reduce their training costs may not be as effective as implementing policies that increase the residual planning horizon by delaying retirement. The distance to retirement, or the horizon effect, is “the key feature to understanding the economics of older worker employment” (Cheron, Hairault and Langot, 2011, p.1478).

While the policy interest in this area has been increasing, at least in Europe and in the OECD, there are only a few economic studies that have investigated the interaction between training and retirement. Some of these studies have examined the effects of training policies (or adult education policies) on the decision to retire, and some other studies have investigated how changes in minimum retirement age have affected training decisions. To the best of our knowledge, no study so far has compared the effects of training and retirement policies on the decisions to retire and train, or has tried to use the overlapping exogenous variation provided by these policies to estimate the causal effect of training on retirement and of the length of the working horizon on training. In this paper, we try to fill this gap.

We first consider the variation generated by multiple reforms in the rules that determine entitlement to pension benefits, resulting in increasingly more stringent eligibility criteria across adjacent cohorts of individuals. The causal effects of changes in the legislation on the propensity to retire represent the first source of variation that we consider. We overlay to such effects the additional variability that results from the existence of various incentives
to training provision, the intensity of which varies over time and across regions for all cohorts of individuals affected by the above mentioned pension reforms. The causal effects of the availability of training incentives on the number of training episodes experienced by individuals represents our second source of variation. By exploiting variability across cohorts and regions we are thus able to study the relationship between the stock of training and the retirement status at a given age, considering exogenously defined groups facing different rules for pension eligibility, and working in environments characterised by non homogeneous training costs.

We investigate the interplay between training provision, on the one hand, and the decision to retire, on the other hand, using a longitudinal sample of Italian males aged between 45 and 56 and born between 1938 and 1952, whom we observe during the 1990s and early 2000s, a period characterized by the implementation of pension reforms affecting minimum early retirement age and by the presence of (mainly EU funded) regional training subsidies to continuous vocational training, introduced in 1994 and resulting in a substantial region by time variation in the resources being tendered to encourage continuous vocational training.

We compare the relative effectiveness of policies which affect eligibility to early retirement and policies which subsidize training costs, and find that delaying early retirement by one year reduces the probability to retire by 8.6% and increases the stock of training by 8.1 to 12.5%, depending on the selected measure of training. We also estimate that increasing (regional) training subsidies by one euro per head has virtually no direct effect on retirement and increases the stock of training by 1.3 to 2%. Ceteris paribus, these estimates indicate that Italian regions would have to spend each about 6 additional real euro per head to attain the same level of training generated by exogenous changes in retirement policies. Since, on average, each region had during the sample period about 3 million individuals aged 15 to 64, we estimate that this is equivalent to close to 240 million real euro, a substantial amount in times of fiscal austerity.

---

2 In this paper we measure training subsidies as real euro per capita, where the relevant population is composed of individuals aged 15 to 64.
This is not to say that increasing minimum early retirement age is costless for public finances: if higher minimum retirement age increases the length of unemployment spells among unemployed older workers or generates additional job losses, the government may have to increase the outlays for unemployment benefits. Our evidence in this respect is somewhat dismissive, as we find that a higher minimum early retirement age has had no statistically significant effect on the risk of unemployment.

We estimate the causal effect of a longer working horizon on training and find that policies which increase this horizon by an additional year raise the training stock by 14.1 to 22.7 percent, substantially more than the close to 1% hike induced by an additional real euro per head spent in training subsidies. We also estimate the causal effect of training on retirement and find evidence of a sizeable and statistically significant negative effect: a 10 percent increase in the stock of training reduces the probability of retirement by 5 to 8 percent, depending on the definition of training being used.

We discuss alternative reasons why training may affect retirement and conclude that additional training increases both employability - by reducing the risk of unemployment - and earnings. We also speculate that training may lead to higher job satisfaction, and that this is an additional mechanism which reduces retirement.

The paper is organized as follows. Section 1 reviews the relevant literature. In Section 2 we discuss the relationship between training and retirement that derives from the economic model presented in the Appendix and introduce our empirical strategy. Section 3 is devoted to institutional details and describes Italian reforms of minimum retirement age as well as the provision of regional training incentives. The data are introduced in Section 4 and results are discussed in Section 5. A brief Section 6 with sensitivity analysis precedes our conclusions.

1. Review of the literature

Human capital theory suggests that early retirement systems have a negative impact on human capital formation and training, because they shorten the period during which the worker and/or the employer can reap the benefits of the investment and recoup the costs (Ben Porath, 1967). The effects of training on the timing of retirement are less clear-cut.
Training is expected to increase earnings and the probability of gainful employment. Increased earnings, however, have theoretically ambiguous implications for the timing of retirement. A higher wage yields greater foregone earnings if the worker retires and a higher lifetime income, which raises the value of leisure and the incentive to retire early.

In 1999 the American Association of Retired Persons reported that 80% of baby boomers were expected to postpone retirement, quoting economic reasons as one of the major motivations for continuing labour force participation. Since the eighties, corporate retrenchment and technological change have put substantial labour market pressures on workers aged 50 years and older by more or less compelling them to stay abreast of new techniques (Farber, 1997). The effects of technological change on the retirement choices of individuals are discussed by Bartel and Sicherman (1993). They suggest that workers employed in industries characterized by high rates of technological change tend to retire later because the net effect of technological change on training is positive. On the other hand, older workers are more likely to retire sooner when an unexpected increase in the rate of technological change occurs.

There are only a few empirical studies investigating the relationship between training and retirement. They focus either on the effects of retirement policies on training, or on the effects of training on the decision to retire. To the best of our knowledge, no study so far has addressed both effects. Montizaan et al. (2010) use a natural experiment in the Dutch public sector to examine the effects of an exogenous increase in expected retirement age on training participation. They find that a shock to pension rights which postponed retirement has had a positive impact on the training participation of older men. However, the effect is rather small.

Stenberg et al. (2012) ask whether adult education, which includes training, delays retirement and increases labour force participation among older Swedish workers and find

---

3 See Leuven and Oosterbeek (2008) for a review of the large literature on the private returns to training, and Behaghel Caroli and Roger (2011) and Picchio and Van Ours (2011) for evidence of the effects of training on the employment of older workers.

4 The positive effect of technological change on the returns to training has to be discounted by the negative effect on individual human capital.

5 Charness and Czaja (2006) argue that older workers are quite capable of learning new skills, albeit at a slower pace than younger workers.
no significant effect on the timing of retirement. In contrast, Kristensen (2012) uses Danish data and find that additional training increases retirement age. The estimated effect, however, is small. A common drawback of the these two studies is that they cannot provide credible causal conclusions as they cannot identify sources of exogenous variation in training.

Drawing on the research by non economists, Simpson, Grelle and Stroh (2002) criticize the labour economics literature for failing to adequately account for trends in adult education. Using the Adult Education (AE) file from the 1995 National Household Education Survey (NCES), they focus on human capital investments by late career workers (50–65) and show that workers aged 50 years and older do not invest in general skills but are more likely to participate in activities that develop focused occupational skills, such as job-related courses and on-the-job computer-based training. However, their study could not establish whether such training is the cause of postponed retirement or if career continuity affords more opportunities for such training.

Our exam of the existing literature suggests that, while there is substantial theoretical work on the economics of training and the economics of retirement, relatively less has been done to model the interaction between training and retirement decisions. This is the gap that we intend to fill in this paper.

2. The relationship between training and retirement

Figure 1 illustrates the decisions to train and to retire, which we derive from optimization principles in the economic model presented in the Appendix. Let $R$ be the probability of retiring; $T$ the training stock, obtained by properly discounting the flow of training investments; $D$ the difference between minimum retirement age, which depends on mandatory rules as well as on endogenous social security contributions, and age; $Q$ the discounted stock of training incentives – provided in Italy by regional governments; $Z$ the minimum early retirement age prescribed by law and modified by pension reforms; $X$ the

---

6Useful reviews of these literatures include Bassanini et al. (2007) and Lumsdaine and Mitchell (1999). See also the option value model by Stock and Wise (1990).

vector of predetermined or exogenous variables, which includes educational attainment and potential experience.

Training $T$ increases with the stock of incentives $Q$ and with the length of the remaining working horizon $D$ and declines when the probability of retirement $R$ is higher – see Eq. (A11) in the Appendix. On the one hand, higher incentives reduce the costs of training. On the other hand, a longer residual working horizon increases training benefits. Conditional on eligibility, a higher probability of retirement reduces the likelihood that future benefits can be reaped, and therefore reduces the incentive to invest.

The retirement probability $R$ declines with distance $D$ – in particular because positive distance precludes eligibility - and is affected by training stock $T$. On the one hand, training can increase wages. As pointed out by the existing literature and discussed in the economic model presented in the Appendix, higher wages have both a substitution effect (toward work) and an income effect (toward more leisure and retirement), with an overall uncertain effect on retirement. On the other hand, training can increase employment and reduce the risk of unemployment by providing marketable skills, especially if these are work related. A lower risk of unemployment is likely to reduce the incentive to leave the labour market for good with an early pension. Last but not least, additional training can lead to better and more satisfying tasks and jobs, and therefore induce employees to retire later.

The Figure contains two exogenous instruments – $Q$ and $Z$ – two closely interacting endogenous variables, $R$ and $T$ – and a “transmission mechanism” from $Z$ to $R$ and $T$ given by distance $D$, which we associate to the minimum early retirement age prescribed by law, $Z$. A simplifying assumption used in the current paper – which does not affect our estimates of the effects of training on retirement and of the horizon effect on training – is that the stock of training affects distance $D$ only marginally, and that this potential additional link can be safely disregarded.

Let the pair $T_0, R_0$ be an equilibrium of the model. We describe the interaction of training and retirement with two linear equations, that can be obtained as first order Taylor expansions of the more general nonlinear relationships discussed in the Appendix:
\( R_a = r_0 + r_1 T_a + r_2 D_a + \eta_R \) \hspace{1cm} (1)

\( T_a = g_0 + g_1 R_a + g_2 D_a + g_3 Q + \eta_T \) \hspace{1cm} (2)

where \( a \) is age, \( \eta \) are random errors and we omit for brevity from the list of covariates both age \( a \) and the vector \( X \) of exogenous or predetermined controls.

In the parlance of the simultaneous equations literature, only equation (1) is identified, because it meets both the order than the rank conditions for identification. Equation (2) fails instead to meet the order conditions, which prescribe that the number of excluded variables be weakly greater than the number of equations minus one. To address this problem, we use (1) into (2) and re-write equation (2) as:

\( T_a = a_0 + a_1 D_a + a_2 Q + \varepsilon_T \) \hspace{1cm} (3)

where \( \varepsilon \) is the error term and the coefficient associated to distance \( D \) measures both the direct and the indirect effects of distance on training. In our study, we can only identify this overall effect.

We model the relationship between distance \( D \) and minimum retirement age as follows

\( D_a = d_0 + d_1 Z + \eta_D \) \hspace{1cm} (4)

Using equation (4) into (3) yields training \( T \) as a reduced form function of the exogenous variables – or instruments - \( Z \) and \( Q \) and of other exogenous or predetermined covariates. We can further substitute (3) and (4) into equation (1) and express the probability of retirement \( R \) as a reduced form function of the two instruments \( Z \) and \( Q \).

Our estimation strategy consists of two steps. First, we estimate the two reduced form equations and equation (4). These estimates allow us to measure the effects of the exogenous policy variables, \( Z \) and \( Q \), on training \( T \), retirement \( R \) and distance \( D \). Second, we estimate the causal effects of training \( T \) on retirement \( R \) – conditional on \( Z \) and \( X \) - and of distance \( D \) on training \( T \) - conditional on \( Q \) and \( X \)(i.e. the vector of predetermined age invariant covariates). Using (4) into (1) we get

\( R_a = \phi_0 + \phi_1 T_a + \phi_2 Z + \varepsilon_R \) \hspace{1cm} (5)
We estimate both (3) and (5) by instrumental variables: in equation (3), we instrument endogenous distance $D$ with $Z$, the minimum early retirement age prescribed by pension reforms, after conditioning for $Q$ and the vector $X$. In equation (5), we instrument endogenous training $T$ with the stock of training incentives $Q$, after conditioning for $Z$ and the vector $X$.

We exploit two overlapping sources of exogenous variation that affect the variables of interest, and combine the information they convey to study the interplay between training provision, on the one hand, and the decision to retire, on the other. The causal effects of changes in the legislation affecting $Z$ on the propensity to retire represent the first source of variation that we consider. We overlay to such effects the additional variability that results from the existence of various incentives to training provision, the intensity of which varies over time and across regions for all cohorts of individuals affected by pension reforms affecting $Z$. The causal effects of the availability of training incentives on the individual stock of training represents our second source of variation. By exploiting variability across cohorts and regions, we are thus able to study the relationship between the stock of training and the retirement status at a given age, considering exogenously defined groups facing different rules for pension eligibility, and working in environments characterised by non homogeneous training costs.

3. Pension Reforms and Training Incentives

In this section we describe the two policy variables $Q$ - the (discounted) stock of real training subsidies tendered by local governments to stimulate the provision of continuous vocational training - and $Z$ - the exogenous set of rules governing early retirement age fixed. These rules are fixed by law and were gradually tightened in Italy starting with the 1995 Dini reform and continuing with the 1997 Prodi reform, resulting in increasingly more stringent eligibility criteria across adjacent cohorts of individuals.

3.1 Tendered training incentives

In Italy, government subsidies to continuing vocational training are managed by regional authorities. Public intervention includes: 1) the European Social Fund (ESF); 2) national
measures (Laws 236/93 and 53/00) and 3) industry based training funds (ITF), managed by the social partners. By and large, these measures are funded by the European Community (Objectives 1 and 3, directives D1 and D2 during the financial period 2000 to 2006 and Objective 4 during the financial period 1994 to 1999) and by a compulsory levy of 0.30% on national payroll (see Appendix 2 in Brunello, Comi and Sonedda, 2012, for further details). We estimate that, during the period 1994-2005, about 3.37 billion euro at constant prices have been tendered by regions to support CVT, with 2.7 billion euro funded by the European Social Fund and the rest from the levy on national payroll.

These resources are transferred from the Community and the national government to regional authorities, which have substantial discretion and autonomy in the management of training funds. For instance, funds received by the national government in a given fiscal year are not necessarily allocated to regional budgets nor tendered within the same period. While some regions have managed to issue invitations to tender a few months after receiving their funds, other regions have either not been able or have decided not to do so. As an example for this, the time lag between the allocation of Law 236 funds from the Ministry of Labour to the regions and the first invitations to tender issued by regions ranged in 2003 from 17 to 484 days.

These funds are typically targeted to firms but also include individual vouchers targeted at employees. While public sector employees can access these funds, the self-employed cannot. Therefore, we exclude the latter for the current study. Table 1 shows for 13 Italian regions and for 2004 both $Q$, the discounted sum of tendered CVT subsidies per head (at constant prices), and the average annual flow for the period 1994-2004. There is substantial variation across regions, with Apulia in the South tendering the least (1.36 euro per head per year) and Emilia in the North planning to spend the most (12.07 euro per head

---

8 Industry based training funds have become operational in Italy from the second half of 2004. Since our sample ends in 2004, we ignore them for the purposes of this paper.
9 Before 1994 there were no training incentives for continuous vocational training.
10 The total number of regions in Italy is 20. This number is reduced in the analysis to 13 by grouping small regions together with neighbouring larger regions.
11 We use a 3% discount rate. See Brunello, Comi and Sonedda (2012) for details.
per year). Not only the level but also the dynamics of the discounted stock of incentives exhibit important regional variation, which does not merely reflect regional trends in productivity. To illustrate, we regress the stock on real regional GDP per head, take the residuals and normalise them to 1 in 1994. Figure 2 plots these residuals - and highlights the difference in regional dynamics.

3.2. Reforms to minimum early retirement age

The Italian retirement system includes both old age and seniority pensions. Until 1992, employees in the private sector qualified for old age pensions at age 60 (55 for females) and for seniority pensions at any age provided that they had accumulated 35 years of social security contributions in the private sector and 25 years in the public sector. Since the minimum age required to be legally employed is 15, until 1992 private sector employees could retire in principle at 50, and public sector employees as early as at 40. Starting with the 1992 reform, eligibility conditions both for old age and for seniority pensions were progressively tightened, with the purpose of containing booming public expenditures. In the case of seniority pensions, to which employees typically become eligible before old age pensions, the minimum number of years of social security contributions remained constant at 35 but the combined early retirement age requirement was progressively increased.

Table 2 shows the changes in eligibility conditions both in the private and in the public sector. Before 1996, estimated minimum retirement age was 50 years in the private sector and 40 years in the public sector. In the private sector, minimum early retirement age was raised to 52 years in 1996 and reached gradually 57 years in 2002; in the public sector, it increased abruptly to 53 years in 1998 (when the so called “baby pensions were finally eliminated”) and increased gradually to 57 years by year 2004, the last year in our data.

---

\[12\] The fact that the stock of tendered training subsidies in 2004 was much higher in Northern Friuli than in Southern Campania does not support the view that subsidies have been targeted at regions with training deficits (typically training is lower in the South).

\[13\] Brunello, Comi and Sonedda (2012) argue that an important source of this variation is the political orientation of regional governments, which have changed on several occasions during the period 1994-2005. In particular, they find that having a government with a centre-left political orientation significantly increases the stock of training incentives.

\[14\] See for instance Battistin, Brugiavini, Rettore and Weber, 2009. In Italy, social security contributions are paid by the employer and the employee.
Define $a_m$ as the minimum age required for early retirement. In Italy this age is obtained by combining two components, one exogenous and one endogenous. The exogenous component is the minimum retirement age prescribed by law, or $a_L$. The endogenous component is the age when an individual has completed the minimum required years of social security contributions (typically 35 years). To illustrate with an example, an individual aged 52 and employed in the private sector in 1992 who had accumulated only 34 years of contributions could not retire before 53. For this individual, $a_L = 52$ and $a_m = 53$. In order to focus on the exogenous variation, we define $Z$ as $Z = \max(a, a_L)$, the maximum between current age and $a_L$. If an individual is older than minimum early retirement age, and therefore not bound by this age, we set his minimum retirement age equal to his current age. If the individual is instead younger than minimum early retirement age, we set his minimum retirement to $a_L$. Conditional on age, an increase in $Z$ affects the retirement decisions of individuals with $a < a_L$. Individuals younger that $a_L$ could only escape the binding constraint if they had accumulated enough years of social security contributions to remain exempt from age limitations (see Table A1 in the Appendix for further details).

4. The Data

We use individual panel data from ILFI (Longitudinal Survey of Italian Households). The ILFI panel consists of five waves, one every two years, starting in 1997 and ending in 2005, and contains information on mobility, education, occupation, training and family resources of a representative sample of Italian households. This dataset is particularly well suited for the purpose of this paper because it has information on retirement decisions, the number and duration of training episodes and several individual characteristics, including education and occupation. In the first wave, the survey collected both current

---

15 This dataset has been used extensively in recent years both by sociologists (see for instance Pisati and Schizzerotto, 2004) and by economists (see Gagliarducci, 2005, Silva, 2007, and Bison, Rettore and Schizzerotto, 2009).

16 Respondents to the survey who left home and formed new households were followed by ILFI across the five waves, and their spouses were interviewed as well.
and retrospective individual information on significant events in a sample of more than 4000 Italian households. In the follow-ups, interviews updated the initial information with additional events and collected all retrospective information for the newly interviewed.

Close to 10,000 individuals were interviewed in each wave. In 1997, households were sampled according to a two-stage stratified procedure. The primary sampling unit was the universe of Italian municipalities in 1996. The secondary sampling unit was the household: within each municipality a random sample of households was extracted and all individuals older than 18 were interviewed. The sample is thus representative of the Italian population at the regional level.

We use the retrospective information on the relevant events which occurred since labour market entry to reconstruct individual labour market histories. We consider as reference cohorts those born between 1938 and 1952, who have been exposed to the overlapping exogenous variation of several pension reforms implemented between 1992 and 1998 and to changes in regional training subsidies introduced in 1994, which produced substantial region by time variation in the resources being tendered to encourage continuous vocational training.

Training in these data includes any program organized by firms, local authorities and industrial associations that takes place after completion of upper secondary education and is not included in vocational tertiary education. Even though the survey is done every two years, we can retrieve for each individual the annual number of training episodes (flows), using the information on the year and the month when each episode was started. We allocate to each year all training episodes that start in the year and add up these flows into the training stock $T$ with the perpetual inventory method, using a 3 percent discount rate.

---

17 When we compare our data with the European Labor Force Survey for the year 2000, we find that in our sample of Italian males aged 45 to 55 84.2% were employed, 3.65% were unemployed and 12.17% were either retired or inactive. In a similar sample drawn from ELFS we find that these percentages were 84.9, 3.1 and 12% respectively.

18 For each year, the individual values of our variables refer to the end of the year. We cannot use the year 2005 because very few individuals were interviewed at the end of the year. Using the available information on the region of residence, we retain in our dataset only those individuals who have not changed region since 1980.

19 See Brunello, Comi and Sonedda (2012) for the method used to estimate the depreciation rate.
For each training episode, we have information on the year and the month when the episode started and was completed, which can be used to compute a measure of duration (in months). We use estimated duration to compute an alternative measure of the training stock, which weights each episode with its duration in months. Both measures represent a substantial improvement with respect to using training incidence, as it is often done in this literature (see Bassanini et al., 2007, for a review)\textsuperscript{20}.

We have information on individual labour market status at the end of each year (employed, unemployed or out of the labour force, retired) and define distance from minimum retirement age $D_a$ as the difference in years between minimum retirement age and age. We calculate the number of years of social security contributions that each individual has accumulated using information on her labour market history, including the spells of inactivity and unemployment. Further details on the method employed is provided in the Appendix. In our sample, a tiny minority of individuals are retired in spite of the fact that their distance from minimum retirement age is negative. This is clearly due to measurement error – either in self-reported retirement status or in estimated distance. Since there are very few individuals in this situation, we have decided to drop them from our final sample.

We avoid having region by year cells with too few observations by aggregating the original 20 regions into 13.\textsuperscript{21} In the aggregated regions, the stock of training incentives – which varies by region and year - is computed as the weighted average of the original regional data, using active population in each region as weight. We illustrate the relationship between training stock and training incentives across regions and time by regressing both variables on regional dummies and trends, regional real GDP per capita and regional unemployment rates and by plotting the residuals in Figure 3. Net of the effects of

\textsuperscript{20} In our sample of male workers aged 45 to 56, the percentage of individuals with no training episode is 69.2; close to 15 percent have had one training episode, 8.8 percent have had between 2 and 3 episodes and 6.9 percent have experienced more than 3 episodes. Of those with at least one training episode (30.8%), 18.2 percent have received at most 12 months on training over their working life, and the rest has had more than one year of training overall.

\textsuperscript{21} The 13 regions are: Piemonte and Val d’Aosta, Lombardia, Trentino Alto Adige and Veneto, Friuli Venezia Giulia, Emilia Romagna, Toscana, Liguria, Marche and Umbria, Lazio, Abruzzo and Molise, Campania, Puglia, Basilicata and Calabria, Sicilia and Sardegna.
macroeconomic and regional variables, there is evidence of a positive correlation between the average stock of training and the stock of training incentives.

We select from the original dataset an initial sub-sample of active individuals aged between 45 and 60, who have been observed at least once during the period 1990 to 2004. More in detail, if an individual is aged 45 in 1990, we follow him until age 59 in 2004. If he is aged 58 in 1990, we observe him from 1980, the initial year in our data, when he was 48, until 1992, when he was 60 years old. We only consider males, who have a much higher activity rate than females: according to the OECD, 89.8% of males aged 45 to 54 and 56.3% of males aged 55 to 59 were active in 2004. In comparison, the activity rate of Italian females was 56% and 30.7%, respectively.

We classify the individuals in the selected sub-sample in three broad cohorts, depending on their year of birth. We assign to the first cohort those born from 1930 to 1937, to the second cohort those born between 1938 and 1952 and to the last cohort those born from 1953 to 1959. We then tabulate the dummy $P$ – equal to 1 if an individual has been exposed to a change in mandated minimum retirement age and to 0 otherwise – by cohort and find that the dummy is always 0 in the former cohort and always 1 in the last cohort. This suggests that for two cohorts out of three the exogenous variation in mandated minimum retirement age cannot be distinguished from a cohort effect. We therefore select for our empirical analysis only the second cohort. This is equivalent to choosing as pivotal cohort those born in 1945 and to select a window of seven years before and after this cohort.

Next, we tabulate the dummy $P$ for the selected cohort by age, and find that it varies within each age group only for those aged 45 to 56 during the sample period. For those aged from 57 to 60, the dummy takes always the value 0, and cannot be distinguished from age effects. Because of this, we remove the older age group from our sample. Our final sample consists of 747 individuals born between 1938 and 1952, and 7613 observations, with an average of 10.24 observations per person.

Table 3 presents the summary statistics for the key variables used in this paper. Average age is 49.97 years, and that the percentage of employed and retired individuals is 85.6%
and 11.6%, respectively. The discounted training stock is equal to 0.348 when measured by counting episodes and to 1.834 when each episode is weighted with its duration (in months). Conditional on having received any training, the discounted stock is equal to 1.329 and to 7.262 respectively. The discounted stock of training incentives is 14.7 real euro per head, and distance from retirement is on average equal to -0.476 years. Only 8.8% of the sampled individuals have more than high school education, and the average age when the first job was started is 18.47.

Figure 4 plots the percentage of retired workers and the average discounted training stock by age. The percentage of retired employees is below 10% at age 52 but rises steeply after this age to reach close to 50% at age 56. The average stock of training increases until age 52 and then drops: workers aged 56 have on average the same training stock as workers aged 47. Figures 5 and 6 plot the percentage of retired workers and the training stock against the gap between age and minimum retirement age (or minus $D$), which is allowed to vary within the window [-5,5]. We notice a clear jump in the percentage retired after the gap becomes positive (distance becomes negative). There is also a downward shift in the training stock at zero gap (or distance), especially in the private sector.

5. Results

As discussed above, our estimation strategy consists of two steps. In the first step, we estimate the reduced form equations, and in the second step we estimate two causal effects: the effect of training on retirement and the “horizon effect”, or the effect of distance (minimum retirement age minus current age) on training. Starting with the first step, we estimate:

$$ Y_t = \alpha_0 + \alpha_1 Q_{it} + \alpha_2 Z_{it} + \alpha_3 X_{it} + \eta_{it} $$

where the outcome $Y$ refers, in the various specifications, to training, retirement status and distance, and we impose $\alpha_i = 0$ in the distance equation. In each regression, the vector $X$ includes regional and age dummies, regional linear trends, the regional unemployment rate

---

22 The majority of the remaining 2.8 percent are unemployed.

23 Notice that distance is negative when age is above minimum retirement age $a_M$. 

and time invariant individual controls, which we treat as predetermined with respect to the training and retirement decisions. These controls are educational attainment (lower secondary, upper secondary and college dummies), the age when the first job was started, and the sector of employment (either public or private) in the early eighties.

The impact of both instruments $Z$ and $Q$ on $Y$ can be interpreted as a differences – in – differences effect. We control for pre-treatment differences in regional trends with linear region specific trends, and for region – specific business cycle shocks – which could affect training and training incentives at the same time – with the regional unemployment rate. We expect that the inclusion of regional dummies, regional trends and the regional unemployment rate substantially alleviate the inference problems generated by the possibility that the region - by - year component of the error term is serially correlated (see Angrist and Pischke, (2009)). We take into account that the policy variable $Q$ is at a different level of aggregation than the dependent variables by clustering standard errors by region and year.

Table 4 presents the estimates of the probability of retirement – using a Probit model, the stock of training – measured in the two alternative ways - as functions of $Z$ and $Q$ and distance as function of $Z$. We find that an increase in mandated minimum retirement age $Z$ reduces distance $D$. Perhaps more interestingly, we also find that exogenous variations in training incentives and minimum retirement age affect positively the individual training stock and negatively the probability of retirement.

When evaluated at its sample mean value, the estimated percentage change in $T$ induced by an additional real euro per head spent in regional training incentives ranges between 1.30 and 1.95, close to 6 times smaller than the effect induced by increasing minimum retirement age by one year (8.1 to 12.5%). Ceteris paribus, this suggests that a policy which provides training incentives to employers and employees can be as effective on training as a policy which increases minimum early retirement age by one additional year if tendered subsidies increase by 6 real euro per head, close to what is actually spent on average per year in the 13 Italian regions. In addition, the estimates in column (1) of the table show that a one - year increase in minimum early retirement age reduces the probability of retirement.
by 8.6%, a much larger effect than the one induced by an additional euro per head spent in training subsidies (-0.34%).

One might hastily conclude from this that raising minimum early retirement age by an additional year – a measure typically motivated by the need to reduce pension outlays rather than by the desire to increase training – can be more effective than direct subsidies at increasing training. Yet delaying retirement may also generate additional outlays of public money, for instance because the unemployment rate of older workers increases, and so does expenditure on unemployment benefits. Whether these additional costs are higher than the estimated training subsidies required to obtain the same effect on training remains an open question.

Table 5 presents the IV estimates of equations (3) and (5). The table is organized in four columns. In the former two, we use the number of training episodes to compute the (discounted) training stock, and in the latter two we weight these episodes with their duration in months. In columns (1) and (3) we report the marginal effect of training on retirement, conditional on $Z$. In columns (2) and (4), we estimate the effect of distance $D$ on $T$ – the horizon effect - conditional on training incentives $Q$. The table also reports the F-statistics of the first stage regressions, which suggest that our instruments are not weak.

We find evidence that a longer horizon before retirement – which corresponds to an increase in $D$ – causes a statistically significant increase in the training stock. The effect is sizeable: we estimate that increasing distance $D$ by one year raises the training stock by 22.7% when we measure $T$ using only training episodes and by 14.08% when we weight each episode with its duration in months. The large estimated horizon effect compounds two effects, the direct effect and the effect operating via a reduction of the probability of retirement. Conditional on distance, marginal changes in the stock of training incentives have small but statistically significant effects on training. We estimate that one additional real euro per head spent in training incentives increase the training stock by 0.79 to 1 percent.

Our estimates also suggest that exogenous variations in the training stock have sizeable and statistically significant negative effects on the probability of retirement. We estimate
that – for a given value of $Z$ - one additional month of training reduces the probability of retiring by 5.23 percentage points. Given that this probability in our sample is 11.6 percent, this is a large effect. This result is confirmed when we measure training with the discounted sum of training episodes. In this case we estimate that a 10% increase in the training stock reduces the probability of retirement by 5.17%. These estimates indicate that the individuals who have been induced to increase their training stock because of the exogenous variations in $Z$ and $Q$ (compliers) have changed substantially their retirement behavior as well.

How do we explain the large negative causal effect of training on the probability of retiring? As discussed in the literature and in the theoretical model presented in the Appendix, the expected effect of training on retirement is ambiguous. We believe that candidate reasons why training matters for retirement should include: a) training encourages employability (see Behaghel Caroli and Roger (2011) for evidence using French data); b) training raises wages; c) training increases job satisfaction by giving access to better and more challenging jobs (see Jones et al (2008)).

In the rest of this section, we use the available data to examine the first two candidates. Because of lack of data on job satisfaction, we are force to leave the exploration of c) to future research. We study the effects of training on employability in our sample by defining the variable $S$, which takes the value 1 for employment, 2 for retirement and 3 for unemployment, and estimate a multinomial logit model, using the same controls as in equation (5) and a control function approach (Imbens and Wooldridge, (2007)), which consists of augmenting the vector of covariates with the residuals from the first stage regression of training on the instruments. Table 6 presents the estimated marginal effects of training (measured as duration) on retirement and unemployment. The first column in the table shows the marginal effects of training and minimum retirement age $Z$ on retirement and confirms the results in Table 5. The second column shows instead that an additional month of training significantly reduces the risk of unemployment. Ceteris paribus, such addition drives this risk to close to zero in the sub-population of affected individuals.

---

24 The error standards are computed by bootstrapping.
Interestingly, we also find that a marginal increase in minimum retirement age increases the risk of unemployment for those who do not change their training stock.

Unfortunately, information on earnings in our sample is available only in four years (1999, 2001, 2003 and 2004). Because we restrict the sample to male workers aged 45 to 56, the sample size is too small and the estimates are rather imprecise. In an effort to improve precision, we resort to an alternative dataset, which has better information on earnings but no information on training. The annual Survey on the Income and Wealth of Italian households (SHIW) by the Bank of Italy runs from the mid 1980s on a broadly biannual basis and include detailed information on (annual) earnings. After pooling data from 1993 to 2004, we have estimated a split-sample IV model (SSIV) by constructing cross-sample fitted values of the stock of training from the first stage estimates based on our main sample and by regressing log real wages on predicted training in the sample of Italian males aged 45 to 55. We find that one additional month of training increases real annual earnings in this sub-sample by 5.7 percent (standard error: 0.022), which corresponds to 1.4 percent per week of training. We tentatively conclude from this evidence that more training leads to less retirement because it reduces unemployment risks and increase earnings, and because the positive substitution effects generated higher earnings prevail on the negative income effects.

6. Sensitivity

In Table 4 we have illustrated the effects of the instruments Z and Q on the training stock and the probability of retirement using the full sample. We allow these effects to vary by education in Table 7 and find that increases in either exogenous minimum retirement age or in training incentives have much larger effects among the better educated than among those with less than upper secondary education. In particular, the impact on training of one additional real euro per head tendered as training incentives is about five times as large among the better educated as among those with less than upper secondary education. See Angrist and Pischke, (2009). Standard errors are obtained by bootstrapping.

25 We restrict our sample to the period 1993-2004 because during this period the data contain a consistent indicator of the sector of employment (private versus public).

26 See Angrist and Pischke, (2009). Standard errors are obtained by bootstrapping.
large for those with at least upper secondary education as it is for those with less education. These results suggest that policy changes affecting either public expenditure for training or minimum retirement age are likely to be much more effective for those individuals who are already less likely to retire early or to train less. To a lesser extent, this contrast emerges also when we compare private and public sector employees, mainly because the latter are more likely to be better educated.

Table 8 presents our IV estimates by education group. There are two main results: first, the horizon effect is larger among the better educated. Second, the causal effect of training on the probability to retire is much larger among the less educated. These estimates suggest two considerations: first, policies which raise training are more likely to payoff in terms of lower retirement for the group who is more likely to retire early (in our sample, the percentage retired among the less educated is above 15%, against 6% for the better educated); second, it is much more difficult to increase training provision among the less educated, at least with the policies considered in this paper.

As an additional sensitivity analysis, we have re-estimated the regressions in Table 4 by allowing the two instruments $Z$ and $Q$ to interact. As shown in Tables 9 and 10, there is some evidence that the two instruments are substitutes in both the retirement and the training equation. These estimates indicate that the effectiveness of higher $Z$ in reducing retirement $R$ is enhanced by higher expenditures in training incentives. On the other hand, they also show that the effectiveness of a higher $Q$ in increasing training $T$ declines when $Z$ is higher. However, the IV estimates obtained by adding as an additional instrument the interaction terms are broadly unaffected with respect to those in Table 5.

Conclusions

Training policies targeted at older workers have been often advocated by policy makers – especially in Europe – to increase labour force participation and reduce pension expenditures. Economists, however, have been sceptical that training subsidies can work, especially for older workers. In this paper, we have estimated the effects of reforms to the eligibility conditions for retirement - which typically lengthen the residual working life of
older workers – and of government training subsidies, which affect the costs of training born by employers and employees, on the training and retirement of older workers.

Our estimates confirm this scepticism by showing that measures which lengthen the working horizon are much more effective on training – and possibly less costly – than measures which try to reduce training costs. Policies which increase minimum early retirement age affect the probability of retirement both directly and indirectly because they increase training, and because higher training leads to higher labour force participation. Using linear probability models for retirement, we estimate that the indirect effect contributes to 20 percent of the overall effect of stricter eligibility conditions on the decision to retire.\(^{27}\)

Policy effectiveness varies also with educational attainment: we have found that policies which increase the working horizon and provide training subsidies are much more effective on workers with at least upper secondary education, who have both a higher training stock and a lower propensity to retire than less educated workers. Unfortunately, both policies are less effective on the broader group of individuals, who make up the bulk of early retirees. We therefore conclude with a rather pessimistic remark. On the one hand, our evidence suggests that increasing training among the less educated payoffs significantly more in terms of lower retirement probabilities. On the other hand, it is much more difficult to raise training in this sub-group of the population, at least with the policies considered in this paper.

---

\(^{27}\) From Eq. (5) we have that 
\[
\frac{\partial R}{\partial Z} = \phi_T \frac{\partial T}{\partial Z} + \phi_R \frac{\partial R}{\partial Z}. 
\]
Using linear probability models we obtain 
\[-0.015 = -0.02 \times 0.149 - 0.012.\]
Appendix

1. The economic model

In this appendix, we develop a discrete time model of training and retirement which motivates Figure 1 in the paper. There are two agents: a senior worker aged \(a\), who cannot retire earlier than minimum retirement age \(a_M\) and later than age \(a_F\), and a firm. Each agent has a single decision, retirement for the worker and training for the employer\(^{28}\). The match between the worker and the firm ends only when the worker retires\(^{29}\). Let \(T_a\) represent the training stock at age \(a\). Conditional on having attained minimum retirement age and on her training stock, the senior worker decides when to retire. Retirement is a permanent decision. The employer decides training intensity at age \(a\), \(T_a\), by comparing the costs and benefits of providing additional training. While the costs are borne at the time of the decision, the benefits span over the worker’s residual working life, and depend on her decision to retire. Each agent in this setup “plays Nash”, and decisions are taken simultaneously: the worker decides when to retire by taking the training decisions of the firm as given, and the firm decides training by taking the retirement decision of the worker as given.

The Retirement decision

Let the benefits from labour market participation at age \(a\) be \(W_a = Y_a(z_a, T_a)\), and the benefits from retirement at the same age be \(R_a = Y_a(z_a, T_a)\). Monetary payoffs to labour market participation vary according to whether the individual is employed or unemployed. Training can affect these payoffs either because it increases the probability of employment or because it increases earnings.\(^{30}\) The utility from labour market participation or retirement at age \(a\) is \(U(W_a)\) and \(T(B_a)\), respectively. The senior worker decides to retire at age \(a\) if \(U(W_a) \geq U(R_a)\) and the following conditions hold.\(^{31}\)

---

28 We choose to model the training decision as an employer’s decision based on the evidence that 80% of vocational training courses are paid for or provided by employers. See Bassanini et al, 2007.

29 We ignore other sources of endogenous and exogenous separations.

30 Leuven and Oosterbeek (2008) and Bassanini et al. (2007) review the empirical evidence on the (private) returns to training. Behagel, Caroli and Roger (2011) and Picchio and Van Ours (2011) show that training positively affect the employment of senior workers.

31 Utility in the final period does not depend on labour market status. Equation (1) uses the simplifying assumption that the effect of employment status at age \(a\) on benefits is small and can be overlooked.
\[ U(W_a) \leq \Gamma(B_a) \quad (A1) \]
\[ \sum_{k=0}^{K} E U(W_{a+k}) \leq \sum_{k=0}^{K} E \Gamma(B_{a+k}) \quad \text{for any } K=1,...,F \quad (A2) \]

where \( E \) is the expectations operator\(^{32} \). \( F = a_e - a \) and we set the discount factor to 1. The following assumption is maintained throughout:

**Assumption 1.** If \( U(W_a) \leq \Gamma(B_a) \), then \( E U(W_{a+x}) \leq E \Gamma(B_{a+x}) \), where \( x > 0 \).

This assumption implies that if current benefits from labour force participation are below the current benefits from retiring, the worker expects this to hold also for future benefits, and ensures that condition \( (A2) \) holds when \( (A1) \) is satisfied. As in Stock and Wise (1990), the worker retires if there is no expected gain from continued labour force participation.

Individual utilities are given by:

\[ U(W_a) = \ln W_a + f + \epsilon_a \quad (A3) \]

where \( f \) is a time invariant unobservable, and:

\[ \Gamma(B_a) = k(a) \ln B_a + f + \eta_a \quad (A4) \]

where \( k(a) \) is the value of leisure, which increases with age and there is:

\[ \epsilon_a = \phi_e \epsilon_{a-1} + \zeta_e, \quad \eta_a = \phi_\eta \eta_{a-1} + \zeta_\eta \]

with \( E \zeta_e = E \zeta_\eta = 0 \).

Thus, the terms \( \epsilon_a \) and \( \eta_a \) are individual–specific random effects that vary with age following a first-order autoregressive process, capturing persistent individual characteristics such as preferences for work versus leisure, health status and unobserved wealth (see Stock and Wise, 1990).

Using \( (A3) \) and \( (A4) \), the probability of retiring at age \( a \) conditional on eligibility is:

\[ \Pr(RT_a = 1 | D_a \leq 0) = \Pr(\epsilon_a - \eta_a \leq k(a) \ln B_a - \ln W_a | D_a \leq 0) \equiv G(a,T_a,D_a) \quad (A5) \]

where \( RT_a \) is a dummy for the retirement status at age \( a \), \( D_a = a_M - a \) is the distance from

\(^{32}\) Utility in the final period does not depend on labour market status.
minimum age, and \( \Pr(R_T = 1 | D_a > 0) = 0 \) because there is no eligibility. When distance is positive, the worker has a certain and positive working horizon before retirement, because the probability of retirement is zero. Upon reaching eligibility for retirement, the worker can retire at any time and the remaining working horizon is uncertain.

The unconditional probability of retiring at age \( a \), \( \Pr(R_T = l) \equiv R_a \), is given by:

\[
\Pr(R_T = l) = R_a = G(a, T_a, D_a) \Pr(D_a \geq 0) \equiv R_a(a, T_a, D_a)
\]

Conditional on eligibility, we expect an increase in age to lead to a higher probability of retirement, because the value of leisure increases faster than the value of work. Conditional on age, an increase in minimum retirement age reduces the probability of retirement because it mechanically rolls back eligibility.\(^\text{33}\) A higher training stock influences the retirement decision because it can affect wages, employment, job satisfaction, benefits and the value of leisure. As shown by Stock and Wise (1990), higher earnings have theoretically ambiguous implications for the timing of retirement, which depend on the shape of the utility function, subjective time preferences, and the link between income from work and pension entitlement.

**The decision to train: employers.**

The (private) employer decides how much to train the senior worker in order to maximize expected profits. In doing so, she takes the retirement decision as given. Let profits net of training costs at age \( a \) be given by:

\[
\pi_a = (1 - \sigma)y(T_a, X) - c(q) \frac{T_a^2}{2}
\]

where \( y \) is output, \( q \) is the flow of training incentives, \( T_a = \tau_a + (1 - \delta)T_{a-1} \) is the training stock and training costs are assumed to be quadratic in the training flow, as in most of the relevant literature\(^\text{34}\).

The firm maximizes:

\[^\text{33}\] Hairault, Langot and Sopraseuth (2010) show that the likelihood of employment is significantly affected by the distance to retirement.

\[^\text{34}\] When the employer belongs to the public sector, we assume that the public agency sets training to maximize the worker's wage \( \sigma y(T_a, X) \) net of training costs. In this case, the optimal choice of training is similar to the one described for the private provider.
The first order condition for a maximum yields

\[ \Omega_a = \sum_{k=0}^{aM-a} \pi_{a+k} + \sum_{k=aM-a}^{aF-a} [1-R_{a+k}] \pi_{a+k} - c(q) \frac{\tau_a^2}{2} \]  
(A8)

\[ \sum_{k=0}^{aM-a} \frac{\partial \pi_{a+k}}{\partial \tau_a} + \sum_{k=aM-a}^{aF-a} [1-R_{a+k}] \frac{\partial \pi_{a+k}}{\partial \tau_a} - \sum_{k=aM-a}^{aF-a} \frac{\partial R_{a+k}}{\partial \tau_a} \pi_{a+k} = c(q) \tau_a \]

A higher minimum retirement age \( a_M \) increases the flow of certain marginal benefits, reduces the flow of uncertain benefits and the flow of benefits accruing from a lower retirement probability if \( \frac{\partial R_{a+k}}{\partial \tau_a} < 0 \). Under the additional simplifying assumptions that retirement occurs when reaching minimum retirement age \( (a_M = a_f) \) and that the marginal profits from training are constant, this condition becomes

\[ D_a \frac{\partial \pi_a}{\partial \tau_a} = c(q) \tau_a \]  
(A9)

Condition (A9) shows that a longer working horizon before minimum retirement age is reached (a higher value of \( D \)) increases the optimal training flow \( \tau_a \).

In general, we expect training investments to increase with the length of the working horizon and with the flow of training subsidies, and to decline when the probability of retiring after reaching minimum retirement age increases. We summarize this with the following implicit function

\[ \tau_a = \tau[q, R_a, a, D_a] \]  
(A10)

As distance \( D_a \) increases, the working horizon lengthens and the incentive to invest in training increases. Starting from equation (A10), we model the training stock as

\[ T_a = T[Q, R_a, a, D_a] \]  
(A11)
where $Q$ is the discounted sum of current and past training incentives$^{35}$.

**The decision to train: individuals**

Whilst the majority of training events is organized and paid by the employer, a minority of training episodes are decided and funded by individuals, who pay the training costs. In this case, the model above needs to be adapted. The key change that we need to make is to assume that decisions are taken sequentially: first, the senior individual decides training, by taking into account both her probability of retirement and the effects that training has on this probability. Second, and conditional on training, the individual decides when to retire by comparing utilities. The sequential structure implies that these utilities are net of training costs, which are bygones when retirement is decided. Under this assumption, individual behaviour can be described with a pair of equations similar to (A6) and (A10).

**Distance from minimum retirement age**

Distance $D_a$ is minimum retirement age $a_M$ minus age and measures the time to and from pension eligibility. This variable captures the horizon effect in our model of training and retirement decisions. By construction, this variable is positive when the horizon is positive and negative when the horizon turns negative, because the individual is older than minimum retirement age. When pension eligibility is reached, the probability of retirement switches from zero to positive and distance $D$ switches from positive to negative.

In Italy, as in most European countries, we distinguish between statutory old age that restricts access to old age pensions and mandatory minimum retirement age, that affects eligibility for seniority pensions. The latter is typically lower than the former. Moreover, access to both types of pension requires that individual age satisfies the eligibility criteria and that a minimum number of years of social security contributions has been completed (35 is the key number for seniority pensions in Italy). While age eligibility rules are exogenous policy measures that cannot be manipulated by individuals, years of social security contributions depend on labour market experience and are therefore the outcome of endogenous individual behaviour. Both eligibility rules

---

$^{35}$ To illustrate why we use the discounted sum of training incentives rather than the simple sum, assume only two periods and let the training flow at time zero and one be given by $t_0 = b_0$ and $t_1 = b_1$, where $b_0$ and $b_1$ are the flows of training subsidies. The training stock at time one is $T_1 = t_1 + (1 - \delta) t_0 = b_1 + (1 - \delta) b_0$, where $\delta$ is the estimated discount rate.
and social security contributions affect $a_M$ and distance $D$. To illustrate, conditional on the same age and eligibility conditions, workers who have started their working life later are likely to have higher minimum retirement age because they have accumulated fewer years of contributions.

Let $Z$ be a function of exogenous mandated rules regulating minimum early retirement age and eligibility for seniority pensions. In Italy, minimum retirement age was tightened by important reforms taking place in the 1990s, the Dini reform in 1995 and the Prodi reform in 1997. We assume that

$$D_a = D[a, Z]$$

(A11)

Distance $D$ is also a function of the vector $X$, which includes education and age when the first job was started.

2. Pension Reforms in Italy

Since 1969 (Law n.153/69) Italy adopted a mandatory PAYG (Pay-As-You-Go) pension system which included both “old age” and “seniority” pensions. The system was quite generous and by the end of the 1980s required urgent reforms to guarantee its sustainability. A stark example of the generosity of the system were the so called “baby pensions”, which allowed married females employed in the public sector to retire and draw generous benefits after having accumulated only 14 years, 6 month and one day of social security contributions. Men in the public sector were slightly less fortunate and could retire after 20 years of contributions.

The 1992 Amato reform (Law n. 503/92) reduced the generosity of benefits by introducing the principle that either 10 years to be applied to workers with more than 15 years of contributions or the entire working life for all the others workers– not just the last five years – should be used to compute average earnings as the denominator of the pension replacement rate. Eligibility to old age pensions was also tightened. The critical age increased gradually for men from 60 years old and 15 years of contributions (55 years old for females) in 1992 to 65 years old and 20 years of contributions (60 years old for females) in 2001. These tighter rules did not apply to workers with at

36 Individuals older than 65 who are not covered by old age pensions receive social pensions.
37 Blondal and Scarpetta, 1998, and Angelini, Brugiavini and Weber, 2009, argue that the generosity of the system has been a key reason for the relatively low labour force participation of individuals aged 55 to 64.
38 The main Italian pension reforms that occurred in 1992, 1995 and 1997 are known as Amato, Dini and Prodi reforms respectively.
least 15 years of accumulated contributions at the end of 1992. The disparity of treatment between older and younger cohorts was maintained in the subsequent reforms, leaving the former relatively unaffected.

Law n. 503/92 also stated the intention to abolish the so called “baby pensions” in the public pensions. Due to heavy resistance, actual implementation of the Law started only in 1998. Law n. 335/95, the so called Dini pension reform changed the system from defined benefits to defined contributions. This epochal reform, however, applied entirely only to the new workers hired after 1995 and did not apply at all to older workers who had at least 18 years of accumulated contributions by the end of 1995.

The Dini pension also changed the minimum age required to access seniority pensions. With the new law, from 1996 employees could retire with 35 years of accumulated contributions only if they satisfied a minimum age requirement (52). This minimum was not binding only for those workers who had accumulated a higher number of years of social security contributions (36 in 1996). Table A1, corresponding to Table B of Law 335/95, describes in more details the eligibility rules for access to seniority pensions.

<table>
<thead>
<tr>
<th>Year</th>
<th>Age and Years of Contribution</th>
<th>Only Years of Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>52 and 35</td>
<td>36</td>
</tr>
<tr>
<td>1997</td>
<td>52 and 35</td>
<td>36</td>
</tr>
<tr>
<td>1998</td>
<td>53 and 35</td>
<td>36</td>
</tr>
<tr>
<td>1999</td>
<td>53 and 35</td>
<td>37</td>
</tr>
<tr>
<td>2000</td>
<td>54 and 35</td>
<td>37</td>
</tr>
<tr>
<td>2001</td>
<td>54 and 35</td>
<td>37</td>
</tr>
<tr>
<td>2002</td>
<td>55 and 35</td>
<td>37</td>
</tr>
<tr>
<td>2003</td>
<td>55 and 35</td>
<td>37</td>
</tr>
<tr>
<td>2004</td>
<td>56 and 35</td>
<td>38</td>
</tr>
</tbody>
</table>

At the end of 1997, the Prime Minister Prodi tightened further eligibility requirements. The rules in Table A1 were maintained for blue collar workers in the private sector and for individuals who had paid at least one year of social security contributions when aged 14 to 19. Tables A2 (corresponding to Table C, Law n.449/97) and A3 (corresponding to Table D, Law n.449/97) illustrate the new eligibility rules after the enactment of Law n.449/97, which introduced different eligibility requirements for the private and the public sector.
2. Estimate of the years of social security contributions

We use retrospective labour market history to compute cumulated social security contributions (in months). There are two ways to accumulate contributions in Italy. The usual way is to be employed or self-employed and to pay contributions. The less usual way is that the government pays contributions. This occurs in the following circumstances:

- Unemployment or similar conditions. The benefits associated to temporary layoffs (Cassa Integrazione Guadagni) and unemployment insurance include the payment of social security contributions by the government ("contributi figurativi");
- compulsory military service;
- maternity leave (since 1994 entitlement to benefits requires at least five years of accumulated contributions);
- sickness leave (up to 12 month before December 1996, 14 months from 1997 to 1999, 16 months from 2000 to 2002 and 18 months thereafter).
Starting in 1996, the contributions paid by the government cannot be counted to attain the 35 years required for early retirement. To estimate accumulated contributions, we add up periods of employment with periods of unemployment, compulsory military service, maternity leave (when applicable) and sickness leave.
Table 1. Tendered regional training incentives. Real Euros per head.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Piemonte</td>
<td>95.23</td>
<td>7.48</td>
</tr>
<tr>
<td>Lombardia</td>
<td>69.45</td>
<td>5.41</td>
</tr>
<tr>
<td>Veneto and Trentino</td>
<td>123.34</td>
<td>9.18</td>
</tr>
<tr>
<td>Friuli</td>
<td>142.34</td>
<td>11.20</td>
</tr>
<tr>
<td>Emilia</td>
<td>151.20</td>
<td>12.07</td>
</tr>
<tr>
<td>Liguria</td>
<td>86.25</td>
<td>6.49</td>
</tr>
<tr>
<td>Toscana</td>
<td>71.56</td>
<td>5.42</td>
</tr>
<tr>
<td>Umbria and Marche</td>
<td>80.97</td>
<td>6.94</td>
</tr>
<tr>
<td>Lazio and Abruzzo</td>
<td>54.69</td>
<td>4.02</td>
</tr>
<tr>
<td>Campania</td>
<td>30.98</td>
<td>2.35</td>
</tr>
<tr>
<td>Puglia</td>
<td>18.39</td>
<td>1.36</td>
</tr>
<tr>
<td>Calabria and Basilicata</td>
<td>29.77</td>
<td>2.02</td>
</tr>
<tr>
<td>Islands</td>
<td>35.08</td>
<td>2.20</td>
</tr>
</tbody>
</table>

Source: IGRUE, Regional Bulletins, Corte dei Conti and ISFOL
<table>
<thead>
<tr>
<th>Year</th>
<th>Minimum retirement age - private sector</th>
<th>Minimum retirement age – public sector</th>
</tr>
</thead>
<tbody>
<tr>
<td>Before 1996</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>1996</td>
<td>52</td>
<td>40</td>
</tr>
<tr>
<td>1997</td>
<td>52</td>
<td>40</td>
</tr>
<tr>
<td>1998</td>
<td>54</td>
<td>53</td>
</tr>
<tr>
<td>1999</td>
<td>55</td>
<td>53</td>
</tr>
<tr>
<td>2000</td>
<td>55</td>
<td>54</td>
</tr>
<tr>
<td>2001</td>
<td>56</td>
<td>55</td>
</tr>
<tr>
<td>2002</td>
<td>57</td>
<td>55</td>
</tr>
<tr>
<td>2003</td>
<td>57</td>
<td>56</td>
</tr>
<tr>
<td>2004</td>
<td>57</td>
<td>57</td>
</tr>
<tr>
<td>Table 3. Summary statistics.</td>
<td>Mean</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>-----------------------------------------------------------------</td>
<td>--------</td>
<td>--------------------</td>
</tr>
<tr>
<td>% retired</td>
<td>11.624</td>
<td>-</td>
</tr>
<tr>
<td>Distance from minimum age (in years)</td>
<td>-0.476</td>
<td>6.578</td>
</tr>
<tr>
<td>Discounted training stock – number of episodes</td>
<td>0.348</td>
<td>0.89</td>
</tr>
<tr>
<td>Discounted training stock – duration in months</td>
<td>1.834</td>
<td>4.52</td>
</tr>
<tr>
<td>Discounted stock of incentives</td>
<td>14.708</td>
<td>25.72</td>
</tr>
<tr>
<td>High school</td>
<td>27.781</td>
<td>-</td>
</tr>
<tr>
<td>College</td>
<td>8.879</td>
<td>-</td>
</tr>
<tr>
<td>Age when first job started</td>
<td>18.471</td>
<td>4.65</td>
</tr>
<tr>
<td>Age</td>
<td>49.974</td>
<td>3.34</td>
</tr>
<tr>
<td>% employed</td>
<td>85.577</td>
<td>-</td>
</tr>
<tr>
<td>% in public sector</td>
<td>40.549</td>
<td>-</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>7613</td>
<td></td>
</tr>
<tr>
<td>Number Individuals</td>
<td>744</td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Reduced form regressions of retirement, distance and training on the instruments.

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>D</td>
<td>T episodes</td>
<td>T duration</td>
</tr>
<tr>
<td>Z</td>
<td>-0.0100***</td>
<td>0.7224***</td>
<td>0.0434***</td>
<td>0.1497***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.031)</td>
<td>(0.010)</td>
<td>(0.044)</td>
</tr>
<tr>
<td>Q</td>
<td>-0.0004***</td>
<td></td>
<td>0.0068***</td>
<td>0.0240***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Percent change in R due to a one-year increase in Z: -8.59
Percent change in R due to one additional euro per head spent in Q: -0.33
Percent change in T due to a one-year increase in Z: 12.44 or 8.15
Percentage change in T due to one additional euro per head spent in Q: 1.95 or 1.30

Observations: 7613
Controls: YES
Clustered SE: YES

Note: Robust standard errors clustered by region and year within parentheses; *** p<0.01, ** p<0.05, * p<0.1. Each regression includes age dummies, linear regional trends, the regional unemployment rate and individual time invariant controls (education, age when the first job started, and sector of employment at entry in the sample. Marginal effects from a Probit model in column (1).
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1) R on T (episodes)</th>
<th>(2) T on D (episodes)</th>
<th>(3) R on T (duration)</th>
<th>(4) T on D (duration)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>-0.1710*** (0.061)</td>
<td>-0.0523*** (0.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>-0.0217*** (0.003)</td>
<td>-0.0246*** (0.005)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td>-0.0749*** (0.016)</td>
<td>0.2584*** (0.077)</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>0.0035*** (0.001)</td>
<td></td>
<td>0.0128** (0.005)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>7613</td>
<td>7613</td>
<td>7613</td>
<td>7613</td>
</tr>
<tr>
<td>F-test first stage</td>
<td>117.8</td>
<td>240.3</td>
<td>240.3</td>
<td>240.3</td>
</tr>
<tr>
<td>Percent change in R due to a 10% increase in T</td>
<td>-5.13</td>
<td>-1.57</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Increase in T
Percent change in T due to a one year increase in D | 21.48 | 14.08 |
Percent change in T due to one additional euro per head spent in Q | 1.01 | 0.70 |

Controls | YES | YES | YES | YES |
Clustered SE | YES | YES | YES | YES |

Note: see Table 4. Marginal effects from a IV Probit model in columns (1) and (3).
Table 6. Multinomial logit regression of the dummy $S$ ($S=1$: employed; $S=2$: retired; $S=3$: unemployed) on $Z$ and $T$ (duration). Marginal effects.

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Retired</td>
<td>Unemployed</td>
</tr>
<tr>
<td>$T$</td>
<td>-0.0522***</td>
<td>-0.0234**</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>$Z$</td>
<td>-0.0295***</td>
<td>0.0047*</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Observations</td>
<td>7613</td>
<td>7613</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td></td>
</tr>
<tr>
<td>Clustered SE</td>
<td>YES</td>
<td></td>
</tr>
</tbody>
</table>

Note: see Table 4.

Table 7. Reduced form regressions of retirement, training and distance on the instruments. By level of education. Training duration.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>Low education</th>
<th>High education</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R$</td>
<td>$T$ duration</td>
</tr>
<tr>
<td>$Z$</td>
<td>-0.0214***</td>
<td>0.0200</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.053)</td>
</tr>
<tr>
<td>$Q$</td>
<td>-0.0007***</td>
<td>0.0125***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Observations</td>
<td>4822</td>
<td>4822</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Clustered SE</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Note: see Table 4.
Table 8. IV estimates. Effects of T on R and of D on T. By education. Using training duration.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(high edu)</th>
<th>(high edu)</th>
<th>(low edu)</th>
<th>(low edu)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R on T</td>
<td>T on D</td>
<td>R on T</td>
<td>T on D</td>
</tr>
<tr>
<td>D</td>
<td>-0.3056***</td>
<td>-0.0379</td>
<td>-0.0067</td>
<td>-0.0793***</td>
</tr>
<tr>
<td></td>
<td>(0.082)</td>
<td>(0.101)</td>
<td>(0.009)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Q</td>
<td>0.0377***</td>
<td>0.0113**</td>
<td>-0.0053</td>
<td>-0.0298***</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.005)</td>
<td>(0.004)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>T</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>2,754</td>
<td>2,791</td>
<td>4,822</td>
<td>4,822</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Clustered SE</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 9. Reduced form regressions of retirement and training on the instruments. With the interaction between Z and Q. Using training duration.

<table>
<thead>
<tr>
<th>Variables</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>-0.0129***</td>
<td>0.1467***</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.043)</td>
</tr>
<tr>
<td>Q</td>
<td>0.0036**</td>
<td>0.2859***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>Z*Q</td>
<td>-0.0001**</td>
<td>-0.0046***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Observations</td>
<td>7613</td>
<td>7613</td>
</tr>
<tr>
<td>Controls</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Clustered SE</td>
<td>YES</td>
<td>YES</td>
</tr>
</tbody>
</table>

Note: Robust standard errors clustered by region and year within parentheses; *** p<0.01, ** p<0.05, * p<0.1. Each regression includes age dummies, linear regional trends, the regional unemployment rate and individual time invariant controls (education, age when the first job started, and sector of employment at entry in the sample.)
Table 10. IV estimates. Effects of T on R and of D on T. Using Z*Q as additional instrument and training duration.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>R on T</th>
<th>T on D</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>-0.2792***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.066)</td>
<td></td>
</tr>
<tr>
<td>Q</td>
<td>0.0117**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td>T</td>
<td>-0.0392</td>
<td>-0.0282***</td>
</tr>
<tr>
<td></td>
<td>(0.091)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

F- test first stage
J - test
Observations 7,613 7,613
Controls YES YES
Clustered SE YES YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Figure 1. A graphical description of the economic model.
Figure 2. Residuals from the regression of the stock of discounted regional training grants per head on real regional GDP per capita. Years 1994-2004. Normalized to 1 in 1994.

Note: PIE: Piemonte; LOM: Lombardia; FVG: Friuli; VEN: Veneto and Trentino; ERO: Emilia; TOS: Toscana; MAR: Marche and Umbria; LAZ: Lazio and Abruzzo; CAM: Campania; PUG: Apulia; CAL: Calabria and Basilicata; ISL: Islands.
Figure 3. Residuals of regressing the (average) training stock (episodes) and training incentives on real GDP per capita, the unemployment rate, regional dummies and linear trends. Italy 1990-2004
Figure 4. Percent retired and the discounted training stock (episodes), by age.
Figure 5. Percent retired, by distance from minimum retirement age.

Retired by sector of activity

- **all**
- **public**
- **private**

- Actual values
- Fitted values
Figure 6. Discounted training stock (episodes), by distance from minimum retirement age.

Training stock by sector of activity

- **all**
- **public**
- **private**

- actual values
- Fitted values
References


Bison, Rettore and Schizzerotto, (2009), The Treu Reform and contractual mobility in Italy. A comparison between labour-market entry cohorts, IRVAPP Progress Report n. 2009-02, Trento


Fleischhauer (2007), The Effects of Pension Reform on Retirement and Human Capital Formation, University of Saint Gallen WP.


Imbens G and Wooldridge P, (2007), Control Function and Related Methods, NBER


Simpson, Greller and Stroh (2002)