Working life and human capital investment: Causal evidence from a pension reform☆

Elisabeth Fürstenau [a], Niklas Gohl [b], Peter Haan [c], Felix Weinhardt [d,*]

[a] *DIW Berlin, Germany*
[b] *DIW Berlin, Potsdam University, Germany*
[c] *DIW Berlin, FU Berlin, Netspar, Germany*
[d] *European University Viadrina, DIW Berlin, CEP/LSE, CESifo, IZA, Germany*

ARTICLE INFO

ABSTRACT

In this paper, we analyze if an increase in the working life leads to more human capital investment via on-the-job training. We obtain RDD-estimates from a sharp date-of-birth cut-off, generated by a pension reform that increased the Early Retirement Age (ERA) by three years for many women in Germany. In our preferred specification, we find that this reform causally increased on-the-job training by 4.4 percentage points – a relative increase of 28.8 percent. We explore heterogeneity and additional outcomes and show that this effect is driven by the behavior of women with high initial levels of education. Our results speak to human capital models as well as policies towards extending or shortening working life.

1. Introduction

Human capital theory, starting with Ben-Porath (1967) and Becker (1962), predicts that the value of human capital investment increases with the payout period of the investment.[1] This important prediction is the basis for explaining the joint increases in life expectancy and educational investments witnessed in most countries starting in the twentieth century; see e.g. Soares (2005), Cervellati and Sunde (2013). At the same time, rising life expectancy pressures pension systems in many developed economies. The standard policy-response has been to increase retirement age in order to increase working life. This reduces the number of recipients and increases the number of tax payers, thus reducing the fiscal burden. While there exists a very rich and still growing literature on effects of pension policy on labor supply, the literature on effects of such policy on human capital investment is scarce.[2]

In this paper, we show that women affected by a public pension reform change their human capital investment by increasing on-the-job training before retirement. This effect is driven by highly educated

individuals and not explained by other dynamic labor supply responses related to the reform. We discuss implications of these findings for pension policy in ageing societies. Moreover, we regard our analysis as providing an empirical test of the key prediction of the human capital theory that the length of the payoff period causally affects human capital investment decisions.

For the identification, we exploit a pension reform in Germany that abolished the earliest retirement option. In Germany, working life is largely determined by state pension rules and the reform that we study effectively removed the early retirement for women at the age of 60. State pension rules not only have economic incentive effects but as documented by Seibold (2021), also in the context of Germany, they set the legal framework and social norm for the duration of working life. We therefore can interpret this pension reform as generating quasi-random variation in the duration of the payoff period, which generates an exogenous increase in the working life.

[1] Learning-by-doing is an alternative explanation for educational investments over the life-course (Killingsworth, 1982; Foster and Rosenzweig, 1995).
[2] We are only aware of three studies: Montizaan et al. (2010), Bauer and Eichenberger (2017), and Brunello and Comi (2015).

We present a simple theoretical human capital model in the Online Appendix to formally derive that individuals indeed have an incentive to increase on-the-job training when their working life is prolonged.[3] The central mechanism for this human capital effect is that the returns to training increase with the remaining working life of an individual. This channel is the same as in models of initial education investment decisions and life expectancy (e.g. Ben-Porath, 1967 and Becker, 1962), where decisions on human capital and working life are taken jointly. This confirms that our setting speaks to human capital theory.

The pension reform we study has two features that make it particularly well suited to provide causal evidence on the effect of working life on human capital investment. First, the pension reform abolished an important early retirement program for women born after 1951 across a sharp cut-off. Women born in 1951 and before could enter retirement at the age of 60 through this pathway. In contrast, for women born in 1952 or later, this pathway was closed; these women can enter retirement only at the age of 63, or later. This means that not only does this reform provide a sharp cut-off, it also provides large variation at the cut-off. In the context of pension reforms, this is an unusual feature, as such reforms are generally phased-in on a (birth)month-to-month basis or provide only smaller variation. As a second key feature, the pension reform was already announced in 1998 and implemented in 1999. Thus, the affected women, i.e. women born in 1952 and aged 47, still had a long remaining working life to benefit from human capital investment.

Our analysis is based on the German Microcensus. This is a representative yearly household survey that covers 1% of all German households (about 370,000 households per year). The data includes information about job-related training, which we use to measure post-schooling human capital investment. The main variable that we use measures the incidence of job-related training in the past twelve month. Importantly, the sample size of this household survey is unusually large, allowing regression discontinuity design (RDD) estimation.

Our main finding is that an increase in the working life causally increases human capital investment: we show that the probability to participate in training increases by about 4.4 percentage points. Depending on the specification, the point estimates correspond to a relative increase of about 28.8%, which suggests that an increase in the working life has sizable effects on training participation. Note that we compute these average treatment effects on the treated (ATT) by quantifying our RDD-estimates in Section 7.[4]

We examine other factors affected by the reform and if these can explain our findings. This additional analysis is important to shed light onto the underlying mechanisms of the training effect. This matters for policy conclusions as well as for the human-capital interpretation. In a recent study, Carta and de Philippis (2021) present evidence that a pension reform in Italy affected labor supply even before retirement age. In our context, we do find some evidence for such forward-looking behavior in labor supply but no significant effect on net income.[5] In more detail, estimates on the probability of employment before the age of 60 are positive, between 2 and 3 percentage points but their statistical significance depends on the degree of the polynomial of the running variable that we include in our estimation. Estimates for unemployment

are negative but never statistically significant at conventional levels. This is consistent with Geyer and Welteke (2021) who study forward-looking effects in our context for women aged 58 and 59 only. They do not find systematic forward-looking employment effects. We believe our results are also consistent with Carta and de Philippis (2021), who study a policy change that shifted normal retirement by seven years in Italy. Thus, it is reasonable that the larger reform in Italy also induced larger and more systematic effects than the reform in Germany, which shifted the early retirement age by three years.

In the context of our analysis, accounting for forward-looking employment effects is of central importance. This is because the Microcensus is a repeated cross-sectional data set and changes in the composition of working women could generate spurious relations between working life and training.[6] It is possible though maybe not plausible, for example, that women who now continue being employed also participate less in training compared to women who remain in employment regardless of the policy. If this was the case, this could generate positive employment estimates in the absence of forward-looking human capital investments. To investigate this empirically, we conduct several additional analyses. First, we estimate effects on training for age-groups and education groups that do not show forward-looking employment effects, and find that the training effects persist and are especially high for college educated women who do not adjust their employment in response to the reform. Second, we show difference-in-discontinuities estimates for employment before retirement, finding no significant effect on employment overall. We demonstrate that while there is some limited evidence for forward-looking behavior for employment outcomes before retirement for subgroups in our setting, these do not interfere with the forward-looking human capital interpretation of our findings on training.

We also conduct a number of additional robustness tests of our main results. First, we use balancing and density checks and show that sorting or jumps in covariates around the cut-off do not drive our results. Second, we conduct placebo analyses to rule out that our results are driven by other aspects than the reform. Next, bandwidth variations around the optimal bandwidth of 12 months as well as donut regressions allow us rule out that the results are driven by bandwidth assumption, outliers, or few observations close to the cut-off. We also present difference-in-discontinuities estimates that estimate RDD effects for affected women against same-age men who were not affected by the reform. This analysis can account for potentially unobserved changes in economic conditions around the cut-off that affect working men and women similarly, and returns similar estimates to our main results.

Testing for heterogeneity, we find that the positive reform effect is driven by individuals who have higher initial education. The pension reform increases training for women with a college degree by more than thirteen percentage points. In contrast, the effect for women without college degree is not significant. Investigating further heterogeneity, we do not find evidence that this positive response is limited to specific firms or regions.

Our study is related to several strands of the literature: First and most directly, to the small existing literature on a link between retirement policy and training. Montizaan et al. (2010) exploit a one-year increase in pension age to estimate effects on training using data on Dutch public-sector workers that was collected one year after the reform. They find evidence for positive effects that is limited to larger firms. Brunello and Comi (2015) exploit the staggered implementation of a series of increases in the minimum retirement age in Italy to estimate training effects for the private sector. Their headline finding is that workers who were hit by a one-year increase in the pension policy in their fifties increased training participation by about seven to eleven

---

[3] There exist several reasons why individuals have a general motivation to invest in training. Most importantly, empirical evidence shows that training has a positive effect on wages and on employment, see e.g. Frazis and Loewenstein (2005), Blundell et al. (2019). Moreover, training can improve the quality of work and can have positive effects on non-pecuniary outcomes (Ruhose et al., 2019).

[4] We show estimates from a range of specifications, most of which are different to zero at the 1% level of statistical significance but of course still estimated with error. We discuss implications regarding (minimum) effect sizes that we can reject in Section 7.

[5] The next chapter of this dissertation explores forward-looking employment effects and underlying driving mechanisms in greater detail.

[6] Note that even with panel data joint outcomes and composition effects remain difficult to disentangle with variation from a single policy.

percent, depending on specification.[7] In contrast to the previous studies we focus on the human capital effect for women rather than men. Since female labor market participation is in most countries considerably lower, it is of particular importance to understand how to increase female employment and human capital accumulation through pension reforms.

Second, we contribute to the literature on effects of pension policy on employment and income. Studies that estimate positive employment effects based on policy reforms include Duggan et al. (2007), Mastrobuoni (2009), Staubli and Zweimueller (2013), Atalay and Barrett (2015), Manoli and Weber (2016), Geyer and Welteke (2021), or based on structural retirement models (Gustman and Steinmeier, 1986; Rust and Phelan, 1997; French, 2005; French and Jones, 2011; Haan and Prowse, 2014). Recently, Carta and de Philippis (2021) also estimate forward-looking employment effects before retirement. Crucially, this literature typically assumes an exogenous process of human capital investment, thus implying that individuals cannot adjust their human capital investment through additional training in response to a pension reform. Notable exceptions are the structural analyses by Fan et al. (2017), Blundell et al. (2019).[8]

Third, our paper is linked to the literature on human capital investment and life-expectancy, surveyed in e.g. Bloom et al. (2019). The causal empirical literature is small and focused on very particular settings. Oster et al. (2013) use variation in life expectancy driven by Huntington disease realizations across individuals, who have ex-ante similar risks. They find effects in line with human capital theory on college attendance and completion, health outcomes, as well as on job training for individuals between the ages of 17 and 35. In developing countries, Jayachandran and Lleras-Muney (2009) use a strong decline in maternal death rates in Sri Lanka and find positive effects on girls' educational investments measured in years of school education and literacy rates. In another important study, Baranov and Kohler (2018) exploit variation in mortality rates related to HIV medication in Malawi to study effects on savings and on children's educational investments. They find positive effects of an increase in life expectancy on both types of outcomes.[9]

2. Institutional setting, data and descriptives

2.1. Pension reform

Before turning to the empirical analysis, we start by summarizing the relevant aspects of the German pension system and the 1999 pension reform that induced exogenous variation in the working life.

The statutory public pension system is the central part of the pension system in Germany. It covers more than 80% of the workforce with the exceptions of groups that are not subject to compulsory pension insurance, most important civil servants, and the self-employed. It includes old-age pensions, disability pensions, and survivor's benefits.

The system is financed by a pay-as-you-go (PAYG) scheme and has a strong contributory link: pension benefits depend on the entire working history. The pension system provides several pathways into early retirement, i.e. claiming retirement benefits before reaching the normal retirement age. In this analysis, we focus on the *pension for women*, which allows drawing benefits starting from age 60.[10]

The 1999 reform abolished the *pension for women* for cohorts born after 1951. Effectively, the reform raised the Early Retirement Age (ERA) for most women from age 60 to age 63, thus prolonging the working life.[11] Women born before 1952 could claim the *pension for women* if they fulfilled certain qualifying conditions. The eligibility criteria were: (i) at least 15 years of pension insurance contributions; and (ii) at least 10 years of pension insurance contributions after the age of 40. According to Geyer and Welteke (2021), about 60% of all women born in 1951 were eligible for the old-age pension for women. In our empirical analysis, we focus only on employed women who are neither self-employed nor civil servants; about 89% of these women fulfill the criteria and, therefore, are eligible for this pathway.[12] The pension reform was implemented when affected women born in 1952 were aged 47. Thus, these women had still a long horizon to benefit from human capital investments.

Geyer and Welteke (2021), Geyer et al. (2018) have evaluated the labor market effects of the pension reform based on administrative data of the public pension insurance accounts and the Microcensus, respectively. The findings of these studies are relevant for the following empirical analysis. Most important the increase in the ERA has a sizable positive effect on the working life on individuals which is the necessary condition for an increase in human capital investment. In more detail, employment rates for eligible women aged between 60 and 62 increase by about 15 percentage points, the combined effect on inactivity and unemployment has with about 12 percentage points a similar size.[13]

2.2. The German microcensus

The Microcensus is an annual, household-based survey with representative information about the population and the labor market in Germany. Participation in the survey is mandatory. It has a sampling fraction of one percent of the German population (about 370,000 households) and constitutes the largest annual household survey in Europe (RDC of the Federal Statistical Office and Statistical Offices of the Laender, 2015).

In the main analysis, we concentrate on employed,[14] women younger than 60 years born in 1951 and 1952, who we observe from 2005 through 2012. For these years, the data include information about the

---

[7] In ongoing work, Bauer and Eichenberger (2017) estimate effects on training from a one-year increase in a sample of Swiss women who were already age 56 when hit be the reform. They find evidence for positive effects similar to Brunello and Comi (2015).

[8] There is empirical evidence that pension reforms have positive effects on life expectancy of men Kuhn et al. (2010), Fitzpatrick and Moore (2018), and Bertoni et al. (2018) find effects on healthy behavior while working for Italian men. However, estimates for effects on life expectancy for women are small and not statistically significant at conventional levels (Fitzpatrick and Moore, 2018).

[9] Note that several studies also discuss the theory of human capital investment through training and provide empirical evidence about the effect on labor market outcomes in form of wages, job security, and employment probability, see, among others, Pischke (2001), Zweimueller and Winter-Ebmer (2000), Barrett and O'Connell (2001), Leuven (2005), Frazis and Loewenstein (2005), Picchio and van Ours (2011), and Ruhose et al. (2019).

[10] In addition early retirement is possible via : (1) the *invalidity pension*; (2) the *pension after unemployment or after old-age part-time work*; and (3) the *pension for the long-term insured*; for more details see Geyer et al. (2018). For a more general description on the German pension system, see Boersch-Supan and Wilke (2004).

[11] The *pension after unemployment or after old-age part-time work* was abolished at the same time as the *pension for women*. However, this does not affect our analysis, as the ERA for this pension type was already 63.

[12] We do not observe eligibility at the individual level and quantify what this implies for the interpretation of the reduced form estimates in Section 7.

[13] In a previous working paper version of this paper, see Gohl et al. (2020), we supplemented our main analysis with data from the German Socioeconomic Panel (SOEP). This longitudinal dataset allows us to observe whether a women who participated in training indeed worked longer due to the pension reform and to scale the training effects by employment effects. In particular, we used the pension reform as a "first stage" providing exogenous variation in employment to then in an instrumental variable approach see whether longer employment actually increases training in earlier periods. The results support the notion that this indeed is the case. Unfortunately, the sample size in the SOEP for the relevant cohorts is too low to produce robust evidence.

[14] Women working in "mini-jobs" are not counted as employed.

month of birth and consistent information about participation in on-the-job training.[15] We observe around 1,250 individuals for each birth month in our sample. Thus, overall, the sample includes information on about 30,000 women born in the two cohorts of interest. The Microcensus includes important socio-demographic variables, such as age, education[16] marital status, household income, and firm size. We consider college education and the geographical "West"-dummy as predetermined and include these as controls. The other variables are potentially endogenous, which we keep in mind when using these for balancing checks.

The Microcensus provides in addition information if an employed person has participated in on-the-job training during the twelve months prior to the survey. The training information includes specifically courses that are related to career development, e.g. to improve management, computer, or rhetoric skills.[17] A further question that is included relates to training that is not job-related. Examples of such training are classes in music, sport and health, cooking, or art that many individuals take in their free time, offered through a network of "Volkshochschulen".

On-the-job training in Germany can take place due to several reasons. According to survey evidence from the German Federal Ministry for Education and Training in 2007 more than 50 percent of all on-the-job training participation was self-motivated by the individual, only 32 percent based on firm requirements and the remainder suggested by a superior (TrainingReport, 2010:p.17). Unfortunately, our dataset does not allow to differentiate between these reasons and thus our training measure includes mandatory as well as voluntary training.[18]

2.3. Descriptive analysis

In Table 1, we provide descriptive information about the key variables in our sample for women born in 1951 and 1952 who we observe over the 2005 to 2012 period in the Microcensus. In addition, we present in the first column the statistics for all employed women born in cohorts 1940–1985 aged 30 to 65. Finally, in the last column we show if the variables significantly differ between the two cohorts we use later for the analysis. Women born in 1952 have a significantly higher level of training participation, are slightly more likely to have a high level of education and are less likely to live in West Germany. Crucially, in order to account for potential confounding effects stemming from these differences we later conduct a balancing test and control for covariates such as the educational level and whether the observed women lives in former West or East Germany.

In Fig. 1, we provide further general evidence about the age and cohort pattern of training for employed women. In Fig. 1, we focus on all employed women born between 1940 and 1985. We find a declining pattern that is explained by cohort and age effects. Training rates are
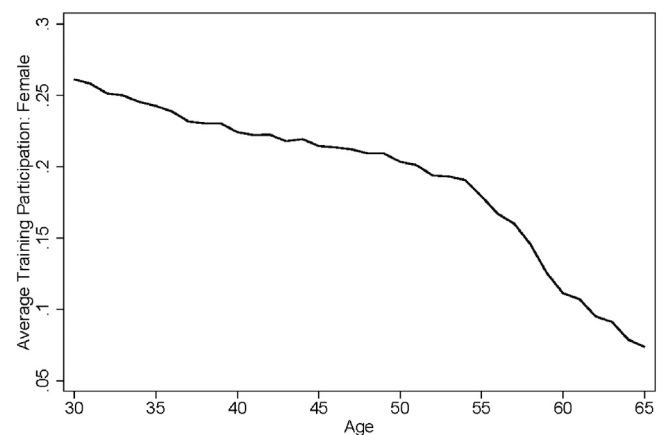

Fig. 1. Average on-the-job training participation by Age.
Notes: Fig. 1 plots the average on-the-job training participation by age for all employed women of the cohorts 1940-1985 using the Microcensus from 2005 onward.

above 25% at the age of 30 and then monotonically decline to about 10% at the age of 60.

This descriptive pattern could already be interpreted as evidence in favor of human capital theory and the theoretical prediction that training should decrease towards the end of working life, but it is important to re-iterate the following: this descriptive figure mixes up cohort and age effects and we know that job-related training is increasingly important (Köller et al., 2017). One result of such unobserved time-trends could be that younger cohorts might have higher training levels throughout their working life and, thus, the age effect could be non-existent or even upward sloping to generate this overall pattern. Thus, descriptive evidence, like that shown in Fig. 1, is relevant for documenting the incidence of training for different age groups but cannot inform causal questions.

To shed some more light on this, Fig. 2 shows age-specific training participation rates for the two cohorts from the Microcensus that we use in the RDD to estimate the training effects. Two observations are of interest: First, the younger cohort of 1952 shows a higher incidence of training for all ages 53 to 59, when compared to the older cohort of 1951. In this Figure, this difference is a combination of the reform effect, since only the younger cohort was affected by the 1999 pension reform, as well as general differences that might occur between cohorts. In our analysis, we control for the latter using the RDD design. Second, the age-trend from 53 to 59 is negative, but only marginally so. This stands in stark contrast to the age-pattern shown in Fig. 1 and underlines the necessity to separate out cohort from age-effects for a causal interpretation.

In Figure A.1 in Online Appendix B, we turn to the training pattern by initial education. We find that the probability to train increases with the level of education. Specifically, employed women with no college degree (ISCED < Lower Tertiary) have training rates of about 11.7 percent. In contrast women with a college degree or higher tertiary education (ISCED ≥ Lower Tertiary) have training rates of 37.1 percent. Again, this figure is merely of descriptive nature, but the differences by educational level motivates our heterogeneity analysis along this dimension.[19]

3. Empirical method: RDD

In the RDD-analysis, we exploit the 1999 pension reform to estimate the effect of an increase in working life on human capital investment.

---

[15] Before 2005, the Microcensus only provides information about the birth year and the definition of training changes at several points in time. Therefore, the extension of the sample before 2005 would require additional assumptions. The Microcensus does not include information about training activities of unemployed, therefore we only consider the effects on employed individuals aged 53–60.

[16] Education is measured with ISCED 1997 levels: based on this information we can differentiate between women without college degree or with college degree or more. In total there are six ISCED levels in this categorization and anyone with a level of five or higher is classified as having a college degree.

[17] The exact wording of the question reads: *Did you, in the last 12 months, take part in any form of vocational training? Examples of vocational training are occupational re-training, courses for career development, and general training courses in, for example, the fields computing, management, and public speaking.*

[18] Note that it is unlikely that there is any difference in the participation in mandatory training between cohorts 1951 and 1952 with regards to this mechanical form of training prior to reaching the age 60, i.e., when individuals in the cohort 1952 are more likely to work longer.

[19] In Online Appendix Figures A.3 and A.2, we also show figures with/without college for the cohorts 1951 and 1952. The patterns are similar to the discussion of Fig. 2 vs. 1 above.

Table 1
Summary statistics: Employed women.
Source: Microcensus 2005–2012, own calculations.

	All cohorts	Cohort 1951	Cohort 1952	Diff. 1951-52
On-the-Job training participation	0.205	0.1543	0.1717	−0.0174***
	(0.403)	(0.3613)	(0.3772)	(0.0044)
Large corporation	0.504	0.5222	0.5247	0.0025
	(0.403)	(0.4995)	(0.4994)	(0.0060)
High level education	0.225	0.1887	0.1990	−0.0103**
	(0.417)	(0.3913)	(0.3993)	(0.0047)
Medium level education	0.651	0.6518	0.6560	−0.0042
	(0.477)	(0.4764)	(0.4750)	(0.0056)
High HH income	0.372	0.3296	0.3380	−0.0084
	(0.483)	(0.4701)	(0.4730)	(0.0056)
Married	0.631	0.7203	0.7165	0.0038
	(0.4001)	(0.4489)	(0.4507)	(0.0038)
West-Germany	0.810	0.7685	0.7582	0.0103**
	(0.0391)	(0.4218)	(0.4282)	(0.0050)
Observations	649,181	13,649	14,870	flushleft

Notes: Table 1 depicts average values of outcome variables and covariates for the female working population for women aged 30-65 (Cohorts 1940–1985) as well as by control (1951) and treatment cohort (1952). Standard deviations for average values are depicted in parentheses. Differences between cohort-specific mean values are tested by a two-sided t-test. Standard errors are depicted in parentheses and significance levels follow ∗ p 0.10, ∗∗ p 0.05, ∗∗∗ p 0.01.
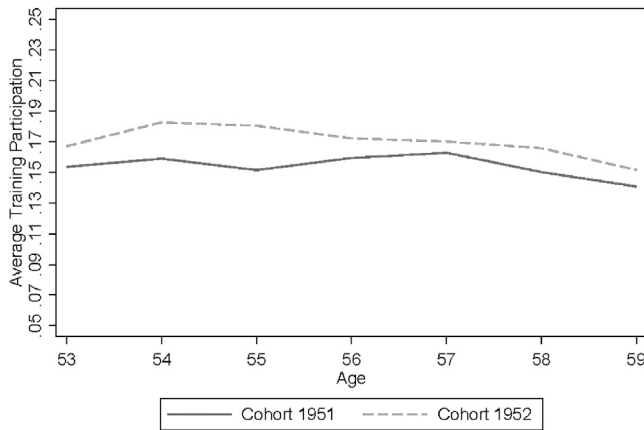
Fig. 2. Average on-the-job training participation for sample group.
Notes: Fig. 2 plots the average on-the-job training participation for all employed women of the cohorts 1951 and 1951 using Microcensus data from 2005 onward.

The reform leads to an arbitrary and distinct cut-off for women born before and after December 31, 1951, which determines assignments into the treatment and the control groups.

More formally, and similar to Geyer and Welteke (2021), in the empirical analysis the standardized woman's month of birth is the running variable M that determines treatment D as one if she was born after December 31, 1951, and zero otherwise:

$$D_i = \begin{cases} 1, & if\ M_i \geq c \\ 0, & if\ M_i < c \end{cases} \quad (1)$$

For identification of a causal effect, it is important that no manipulation of the month of birth for women born in 1951 and 1952 and no selection into or out of treatment is possible. As a result, the treatment and control groups should be otherwise comparable around the cut-off. We provide supporting evidence based on balancing tests of important pre-policy covariates of the 1951 and 1952 birth cohorts as well as by moving the cut-off to hypothetical placebo dates. Moreover, as discussed e.g. in Geyer and Welteke (2021), no other relevant policy reform differently affected women born in 1951 and 1952.[20]

In the main specification, we implement the RDD in the following regression model:

$$y_i = \alpha + \beta D_i + \gamma_0(M_i - c) + \gamma_1 D_i(M_i - c) + X_i\delta + \sum_a \kappa_a A_{it} + \sum_y \lambda_y Y_y + \varepsilon_i \quad (2)$$

D_i is a dummy specifying treatment, that is equal to 1 if a woman is born 1.1.1952 or later and 0 otherwise. A woman's date of birth, measured in months, is described by M_i and c is the cut-off date for the increase in Early Retirement Age, ERA (January 1, 1952). The difference between a woman's birth month and the beginning of the ERA increase, $M_i - c$, gives the running variable. The running variable across all specifications is interacted with the treatment variable D_i to allow for different slopes before and after the cut-off. γ_0 is the coefficient of the running variable and γ_1 is the coefficient of the interaction term. In addition, we account for further explanatory variables X_i, including predetermined education and regional information. More precisely, we control for high/college education (ISCED 1997 ≥ Lower Tertiary), medium education (ISCED 1997 categories 3 and 4) and an indicator variable for west Germany. We also include age fixed effects $\sum_a \kappa_a A_{it}$ and survey year fixed effects $\sum_y \lambda_y Y_y$ to control for potential difference in the survey years and differences in the macroeconomic environment.

In our main approach, we estimate this specification using local regressions and include polynomials up to the second degree in the running variable $M_i - c$ and its interaction with the treatment indicator $D(M_i - c)$.[21] Moreover, we estimate OLS regressions (linear and quadratic) and test for robustness for various bandwidth choices. The outcome variable Y in our analysis is on-the-job training, which is dichotomous i.e. taking on the value 1 if a women has participated in training in the last twelve months and 0 if she has not.[22]

For some robustness checks, we also implement a difference-in-discontinuities approach. The difference-in-discontinuities approach in the setting of this paper effectively estimates the difference between cohorts 1951 and 1952 for women and men and compares them to each other. Empirically, it is implemented by interacting the regression equation for the RDD with an indicator function, equal to one for the actually treated sample, i.e. in this case women. F_i thus is a dichotomous variable equaling one if the individual is female. The equation below details the estimation approach. Note that all variables

[20] Note that in 2007, there was another pension reform, which shifted the normal retirement for the 1952 cohort by an additional month compared with the 1951 cohort.

[21] Local polynomials are estimated using the Stata package "rdrobust" (Calonico et al., 2018a).

[22] Estimation results based on a probit model (not reported) show very similar results.
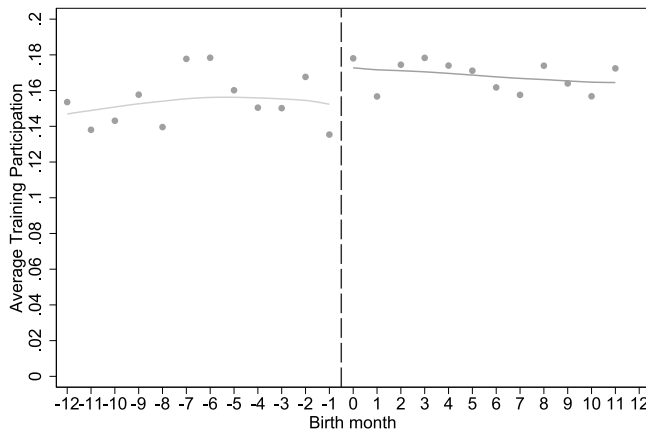
Fig. 3. On-the-job training around the cut-off date.
Notes: This figure plots average on-the-job training over the running variable, which is the distance to the cut-off date measured in birth month (x-axis). The fitted lines are local linear regressions using a first degree polynomial, a triangular kernel. In total, there are observations for 13,658 individuals below the threshold and 14,873 observations for individuals above the threshold. Source: Microcensus, own calculations.

follow the same notation as in the main RDD approach.

$$y_i = \alpha + \beta F_i D_i + \gamma_0 F_i (M_i - c) + \gamma_1 F_i D_i (M_i - c) + \eta_0 F_i + \eta_1 D_i \qquad (3)$$
$$+ \eta_2 (M_i - c) + \eta_3 D_i (M_i - c) + X_i \delta + \sum_a \kappa_a A_{it} + \sum_y \lambda_y Y_y + \varepsilon_i$$

By implementing this strategy we explicitly control for any macro (economic) developments over the estimation period by using men as a control group. This is possible as the end of the *pension for women* did not impact men directly. Further, by including men as an additional control group we can explicitly rule out turn-of-the-year and seasonality effects, i.e. effects caused by the differences in birth years and/or months between control groups and treatment group.

4. The reform effect: Main results

4.1. Training effects: Graphical analysis

Fig. 3 shows participation rates in training by month of birth, 12 months before and after the cut-off birth date, 1.1.1952, for all employed women in their later working life, i.e., when they are aged between 53 and 60. The share of employed women participating in training is higher after the cut-off. Specifically, the average rate of participation among those born in the 12 months before the cut-off is approximately 15.4%. After the cut-off date, the graphs show a jump in the average rate of training participation for employed women under 60 to more than 16.5%.

Importantly, and in contrast to the descriptive evidence discussed in Section 2.2, women close to the cut-off are of almost identical age, thus cohort-effects are an unlikely explanation for this jump. While graphical RDD-evidence can be informative, eye-balling alone can be misleading. Thus, in the next section, we examine the robustness and significance of the graphical evidence using various choices in the RDD framework.

4.2. Training effects: RDD results

To quantify the effect of an increase in the working life on the investment into human capital, we use the RDD described in Section 3. In Table 2 we present the estimation results for different specifications with observations 12 months before and after the cut-off date (we discuss alternative bandwidth choices in Section 5.4). The first two columns present results from regressions estimated using OLS with

Table 2
Regression discontinuity: Training effect.
Source: Microcensus 2005–2012, own calculations.

	OLS		Local regression	
	(1)	(2)	(3)	(4)
Without covariates				
Treatment variable	0.0157	0.0352***	0.0235**	0.0418***
	(0.0147)	(0.0142)	(0.0120)	(0.0131)
With covariates				
Treatment variable	0.0158*	0.0360***	0.0272***	0.0395***
	(0.0077)	(0.0096)	(0.0079)	(0.0088)
Age FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Running variable	Linear	Quadratic	Linear	Quadratic

Notes: Standard Errors in parentheses; clustered at birth month level. Sample includes twelve month before and after reform cutoff. Significance levels: * p 0.10, ** p 0.05, *** p 0.01; Pre-Policy Mean: 15.43 percent; Number of observations 28,519. Specifications including covariates control for a high initial level of education indicator (ISCED 97 Level 5-6), and a variable for medium education (ISCED 97 Level 3-4) as well as an indicator variable for west Germany.

polynomials with linear and quadratic specifications of the running variable. The two remaining columns show the corresponding local estimates. Moreover, the table includes regressions without and with additional control variables. Age and year fixed effects are always included. Standard errors are clustered at the standardized birth month level and are reported in brackets. Our inference is robust to a specification without clustered standard errors using robust standard errors as suggested by Kolesar and Rothe (2018).

The results of these different specifications all point in the same direction, despite some differences in the magnitude of the point estimates: the increase in the Early Retirement Age has a positive and significant effect on the investment in training. Although positive, the linear OLS-specification in the top panel (without covariates) in Column 1 is not statistically significant. However, since this specification does not control for non-linear patterns around the cut-off, it is *a priori* not the preferred specification. In contrast, the local regressions in Table 2 consistently show positive and significant estimates in similar magnitude across all specifications. Following Cattaneo and Titiunik (2022), we focus on the local regressions as our main approach and only report these results in the remaining tables in the main text. In particular, we treat the quadratic local regression as our preferred specification, as this allows to control for potential non-linear patterns around the threshold.

The point estimate in the local regression approaches show that the probability to participate in training increases between 2.4–4.2 percentage points. This is a sizable effect that would translate into a relative increase of about 15.6–27.3% given the pre-reform share in training of 15.4%. Our preferred estimate in column (4), the local regression with quadratic polynomial and with controls, estimates a training effect of four percent, equivalent to a 26% increase in training. As discussed in Section 2, these estimates present intent-to-treat (ITT) estimates because not all women were affected by the reform. We return to the quantification of these ITT-estimates, when we derive estimates for the corresponding average treatment effects on the treated (ATT) in Section 7.

4.3. Training effects: RDD results, by age

We now go into more detail and estimate the reform effects by age. We have already discussed the training-age profile descriptively in Section 2.3. Theoretically, we would expect that relatively younger women affected by the reform have a higher incentive to invest into more training. On the other hand, later retirement induced by the reform might be less salient for individuals further away from retirement age. To investigate age-heterogeneity empirically, we focus on the quadratic specification with covariates.

Table 3
Employment and income effects.
Source: Microcensus 2005–2012, own calculations.

Outcome Variable:	(1) Emp.	(2) Emp.	(3) Unemp.	(4) Unemp.	(5) Income	(6) Income
Without covariates						
Treatment variable	0.0225**	0.0143	−0.0104	0.0099	−1.9007	4.2213
	(0.0092)	(0.0189)	(0.0109)	(0.0121)	(17.9040)	(21.4920)
With covariates						
Treatment variable	0.0249***	0.0135	−0.0130	0.0103	−4.1830	−4.0053
	(0.0091)	(0.0110)	(0.01085)	(0.0103)	(15.2000)	(16.8490)
Age FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Running	Linear	Quadratic	Linear	Quadratic	Linear	Quadratic
N	54,087	54,087	54,087	54,087	50,035	50,035

Notes: Table 3 shows results from local linear and quadratic RDDs for a range of further outcome variables, which are denoted at the top of the table. The variable income measures the personal income denoted in Euro. All specifications control for a high initial level of education indicator (ISCED 97 Level 5-6), and a variable for medium education (ISCED 97 Level 3-4) as well as an indicator variable for west Germany. Standard Errors in parentheses; clustered at birth month level. Sample includes twelve month before and after reform cut-off. Significance levels: * p 0.10, ** p 0.05, *** p 0.01. Source: Microcensus 2005-2012, own calculations.

In Figure A.5 in the Online Appendix we present the results. Overall, we do not find a clear age gradient. The point estimates for the different ages are quite similar. Given the smaller sample size in the age-specific regressions, the confidence intervals are relatively large, thus not all age-specific effects are significant.

One explanation for this finding is the salience of the reform effect. On the one hand, from a human capital perspective training should be most benefiting at younger ages, so the reform effect stronger initially. On the other hand, training in response to extension of the working horizon can be more salient the closer an individual is to potential retirement. Salience could therefore increase with age, thus contributing to the age-pattern documented here.

4.4. Forward-looking employment and income effects

Employment. We now test for direct effects on employment of the reform for pre-treatment ages, i.e. before the earliest retirement age of 60. This is of particular relevance because the Microcensus data set is a repeated cross-section and we base our analysis sample on women in employment before the age of 60. Thus, any effects of the reform on employment could induce sample selection and bias our estimates - or more precisely: the human-capital interpretation of our estimates.

The results of this additional analysis are presented in Table 3. Here, we show estimates similar to Specification 2 using linear and quadratic local regression specifications but for the population of all women in these age groups that responded to the Microcensus survey. In columns (1)–(4) employment and unemployment levels before the age of 60 are used as outcome variables, each coded with a dummy variable that equals one if the person is employed/unemployed at the time of the survey. The results show that there are no significant effects on unemployment, i.e., see columns (3)–(4). The effects on employment are mixed. Employment effects, depicted in columns (1)–(2) are positive in the linear local specification. These significant effects in the employment outcome need to be examined more carefully. We therefore conduct additional tests in order to rule out that potential employment effects induced by the reform drive our results.

First, we provide results by age groups in order to see which ages drive the employment estimates. Figure A.6 in the Online Appendix presents the results for employment. The visual examination across Figure A.5 (effects on training, discussed above) and Figure A.6 (effects on employment) shows that training and employment effects take place at different ages. More precisely, the employment effects are driven by age 55 and to a smaller extent by age groups 58 and 59.[23] In contrast,

the training effects are driven by age groups 53, 56, 57, 58. In a second step, we exclude all age groups for which we find significant employment effects and repeat our main estimation for the training outcome.[24] Table A.4 in the Online Appendix shows the results. The point estimates in both the linear and the quadratic specification are somewhat larger than in the estimation for the whole sample. Further, the training estimates remain highly significant. If employment effects were driving the training estimates, we would have expected smaller and insignificant effects.

Taken together, there is some, albeit not very strong, evidence for forward-looking employment effects in our setting. However, these occur largely for different age-groups of people and therefore do not mechanically drive our findings on training. This confirms our interpretation of the training estimates as evidence for forward-looking human capital investment.

Income:. Another outcome that could be affected by the reform is income. This is because existing empirical evidence shows that training has a positive effect on wages (Frazis and Loewenstein, 2005; Blundell et al., 2019). Moreover, training can improve the quality of work and can have positive effects on non-pecuniary outcomes (Ruhose et al., 2019). The Microcensus does not include information on *gross* labor earnings but on *net* income on the individual and the household level. Net-income includes in addition to labor earnings information on transfer and tax payments. The results shown in Columns (5) and (6) of Table 3 show that there is no significant effect on income in either the linear or quadratic local specification.

In Table A.1 in the Online Appendix we further repeat our estimations for a sample of employed women only. Panel A depicts the results for all employed women and Panel B for all women. Across all specifications, none of the coefficients is statistically significant at conventional levels.

This null effect is not surprising. First, we do not find strong effects on employment behavior and second the net-income measure includes as well transfer income.[25]

5. Robustness

This section is dedicated to providing additional robustness tests, supporting our main results. First, we conduct balancing tests, which

---

[23] The point estimates are borderline significant at the ten percent level in the preferred quadratic specification. When using a linear specification of the running variable the results are not significantly different from zero.

[24] When using employment as an outcome and excluding the described age groups, we no longer find significant employment effects.

[25] The SOEP, another German household level dataset, contains better measures of income but cannot be applied to the RDD-analysis due to smaller sample size. A DiD analysis based on the SOEP does not produce conclusive results and is included in a working-paper version of this article (Gohl et al., 2020).

show that no other factor potential influencing training varies discontinuously around the cut-off. Second, we provide a density test to assess whether sorting patterns around the cut-off may undermine our identification strategy. Next, we conduct placebo tests shifting the cut-off date and focusing on private training, which following human capital theory should not be affected by the reform. Further, we validate our RDD identification approach using difference-in-discontinuities regressions. Last, we provide results for different bandwidths and conduct donut regressions for our RDD approach. This allows to check the sensitivity of our approach to bandwidth choices and outliers.

5.1. Balancing and density tests

Balancing. The assumption underlying the RD design is that other factors vary smoothly across the cut-off. We provide support for this assumption by using individual control variables as outcomes using the same specification as in Table 2. The resulting estimates are presented in Table 4.[26] The Microcensus does not offer many variables that can be safely considered as pre-determined with respect to the 1999-policy change. Therefore, in this first balancing analysis we restrict our attention to whether the individual has a college education and an indicator for "West" (Panel A). We aggregate the ISCED educational levels into two education groups of women with "college" and "non-college". The latter "West" indicator is a dummy variable that equals 1 if the individual was in West Germany in 1989. For both balancing variables and across all specifications, this analysis reveals no jump at the threshold. This supports the underlying RDD assumption of no other changes across the threshold. In addition, we present the balancing of further variables that potentially might be affected by the pension reform. In Panel B, we show that the reform had no significant effect on sorting into big companies, or marital status.[27]

Next, in an additional analysis we seek to rule out more thoroughly that differences in the composition between the cohorts are driving our results. In particular, Figure A.1 in the Online Appendix, clearly shows that initial education and on-the-job training are strongly correlated. We thus use each ISCED level as a separate outcome indicator variable, repeating our balancing exercise for these outcomes. The results from using the local quadratic polynomial specification are shown in Figure A.4 in the Online Appendix. Overall, for most categories both cohorts are balanced. For categories *post-secondary* and *higher tertiary*, we find negative and significant point estimates, suggesting that individuals in cohorts 1952 are less likely to be in these categories. In order to rule out that such compositional differences drive our results, we therefore repeat our main estimations controlling for a more granular set of educational controls by including ISCED level fixed effects.

Table A.2 in the Online Appendix shows that our main RDD estimate, regardless of the specification, does not change in response to the inclusion of additional education fixed effects. Columns (1)-(2), for reference, show the results for the main estimates using local regressions with linear and quadratic polynomials, i.e., the estimate contained in the paper so far. In columns (3)-(4) we replace our education controls[28] with ISCED 1997 level specific fixed effects, i.e., dummy variables for each ISCED Level with the first one serving as a baseline. The point estimates remain highly significant and do not change by much in both the linear and quadratic specification. In a next step, we control for even more granular fixed effects, including subcategories within each ISCED 1997 level as additional indicator variables.[29] Again the main

estimates remain virtually unchanged, which leaves us confident that our estimations are robust to a more comprehensive set of educational controls.

Density. In a next step, we check for potential manipulation around the cut-off. To do so, we implement a density discontinuity test described in more detail by Calonico et al. (2018b). Manipulation of the birth date cut-off and consequent sorting in our setting is difficult and highly unlikely. However, as shown in the previous section there is some evidence for potential sorting into employment (before the age of 60) in response to the reform. This sorting would also show up in a density test of our main sample, as we focus exclusively on employed women in our main specification. As a result, in our setting the usual rule that violations of the density test invalidate the empirical analysis does not readily apply. Instead, this rule only holds for samples that do not show an employment response.

To test this, Table A.3 in the Online Appendix, depicts the test statistics and p-values of the density test using a bandwidth of 12 months as in our main specification and Figure A.9 in the Online Appendix shows the corresponding plot. Column (1) shows the results for our whole sample. In line with the positive employment for subgroups effects the test statistic is significant. In column (2), we only include age groups for which we found insignificant employment effects to account for the mechanical shift in the density. As expected for this sample all p-values of the test now lie above the 10 percent level of statistical significance.

Next, we split the sample into those with a college degree and those without a college degree. Below in Section 6.1, we show that it is exclusively individuals with a college degree who drive the training results, while employment effects for this group are statistically insignificant at conventional levels across all specifications. The density test clearly shows that for individuals with a college degree, i.e. the group driving the training results, there is no significant jump in the density. In contrast, for individuals with no college degree, the density test statistic is highly significant. Overall, the density results show that for the subsamples that drive the training results but not our employment results, there is no significant jump in the density around the cut-off.

5.2. Placebo analysis — different cohorts and private training

We conduct placebo analyses that we present in Table 5. In the first placebo analysis, we artificially shift the cut-off date by one year to 1.1.1950 (Panel A) or to 1.1.1952 (Panel B). Importantly, the pension rules are identical before and after the chosen placebo cut-offs. The shift of one entire year (in either direction) is of particular relevance as this could capture seasonal effects related to the December to January timing of the reform introduction. The result from this additional analysis supports our identification strategy: the treatment effect is very close to zero and not significant in both placebo specifications, with and without additional explanatory variables. Moreover, these effects are precisely estimated and clearly differ from our main findings in Table 2.

Next, we exploit additional information on "private training" that is recorded by the Microcensus. In contrast to job-related training, such "private training" has benefits that also lasts beyond the working life and, thus, we do not expect to find differential take-up depending on the pension rule.[30] The results are presented in Panel C of Table 5. As expected, none of the estimates are statistically significant. Moreover, all estimates are very close to zero and negative, clearly different to the main results found for job-related training.

---

[26] Results also for OLS specifications are similarly balanced (not shown).

[27] The indicator variable *Single* is significant at the 10 per cent level in the local linear regression, however, insignificant across all other specifications.

[28] So far we included a dummy for higher education (ISCED category lower tertiary and higher tertiary) and a medium education indicator (upper secondary and post secondary)

[29] For example, the ISCED level *tertiary education* can be split into vocational and academic education

---

[30] In the Microcensus, private training is classified as general training measures with a predominant private focus to advance one's own skills and knowledge. Examples given for private training in the Microcensus questionnaire are training in the fields of music, sport and health, cooking, or art.

Table 4
Balancing.
Source: Microcensus 2005–2012, own calculations.

Local regressions	(1)	(2)
Panel A: Pre-determined covariates		
College education		
Treatment variable	0.0029	0.0191
	(0.0230)	(0.0307)
West		
Treatment variable	0.01162	−0.0010
	(0.0233)	(0.0326)
Panel B: Further variables		
Big company		
Treatment variable	−0.00900	0.0023
	(0.0080)	(0.0104)
Single		
Treatment variable	−0.0089*	−0.0050
	(0.0051)	(0.0063)
Age FE	✓	✓
Year FE	✓	✓
Running variable	Linear	Quadratic

Notes: Standard Errors in parentheses; clustered at the birth month level. Sample includes twelve month before and after reform cutoff. Significance levels: ∗ p 0.10, ∗∗ p 0.05, ∗∗∗ p 0.01; Outcome variables depicted in bold. Number of observations in descending order of rows: 28,519, 28,519, 27,179, 28,519, 28,175, 28,519.

Table 5
Placebo analysis: Training effects.
Source: Microcensus 2005–2012, own calculations.

Local regressions	(1)	(2)
Panel A: Placebo 1950/51		
Without covariates		
Treatment variable	0.0017	0.0055
	(0.0102)	(0.0153)
With covariates		
Treatment variable	0.0044	0.0035
	(0.0207)	(0.0325)
Panel B: Placebo 1952/53		
Without covariates		
Treatment variable	0.0043	0.0035
	(0.0208)	(0.0326)
With covariates		
Treatment variable	−0.0009	−0.0029
	(0.0086)	(0.0110)
Panel C: Private training		
Without covariates		
Treatment variable	−0.0020	−0.0007
	(0.0030)	(0.0038)
With covariates		
Treatment variable	−0.0021	−0.0009
	(0.0030)	(0.0038)
Panel D: Men		
Without covariates		
Treatment variable	0.0042	0.0063
	(0.0063)	(0.0086)
With covariates		
Treatment variable	0.0059	0.0097
	(0.0062)	(0.0073)
Age FE	✓	✓
Year FE	✓	✓
Running variable	Linear	Quadratic

Notes: Standard Errors in parentheses; clustered at birth month level. Sample includes twelve month before and after reform cutoff. Significance levels: ∗ p 0.10, ∗∗ p 0.05, ∗∗∗ p 0.01; Number of observations Panel A: 24,623 ; Number of observations Panel B: 34,150; Panel C: 28,519, Panel D: 31,557. Covariates: high initial level of education indicator (ISCED 97 Level 5-6), and a variable for medium education (ISCED 97 Level 3-4) as well as an indicator variable for west Germany.

Table 6
Difference-in-Discontinuities: Training effect.
Source: Microcensus 2005–2012, own calculations.

Local regressions	(1)	(2)
Without covariates		
Treatment variable	0.0242**	0.0305***
	(0.0107)	(0.0111)
With covariates		
Treatment variable	0.0212***	0.0266***
	(0.0066)	(0.0065)
Age FE	✓	✓
Year FE	✓	✓
Running variable	Linear	Quadratic

Notes: Standard Errors in parentheses; clustered at birth month level. Sample includes twelve month before and after reform cutoff. Significance levels: ∗ p 0.10, ∗∗ p 0.05, ∗∗∗ p 0.01; Number of observations: 60,076. Covariates: high initial level of education indicator (ISCED 97 Level 5-6), and a variable for medium education (ISCED 97 Level 3-4) as well as an indicator variable for west Germany.

to the main results (i.e. 0.0266 vs. 0.0395, with quadratic polynomial and controls) but this difference is not significant at conventional levels. In contrast, all estimates remain distinct from zero and the null hypothesis of no effect is rejected at high levels of statistical significance. Overall, these results therefore confirm our previous estimates. Note we will refer back to this set of slightly smaller estimates when quantifying effects sizes in Section 7.

We also estimate forward-looking employment effects using the difference-in-discontinuities design (reported in Online Appendix Table A.5). Here, we do not find any significant effects on employment induced by the reform. The fact that the results on training hold in the diff-in-disc setting, but not for the case of employment, supports the earlier conclusion that sample selection through effects on employment before formal retirement age do not drive our findings on training.

5.4. Bandwidth and donut RDD

Bandwidth choices can affect RD estimates, so we carefully examine if and how our main results are sensitive to these.

In Online Appendix Table A.6 we show that our results also hold for additional bandwidth choices of 6, 9, and 12 months. The bandwidth of 12 months is the chosen bandwidth of the endogenous bandwidth selection routine "rdbwselect" using the mean squared error criterion, a triangular kernel and the preferred quadratic polynomial (Calonico

5.3. Difference-in-discontinuities analysis

Table 6 shows the estimates from the difference-in-discontinuities approach outlined in Section 3. The results support the main findings presented above. The results are a little bit smaller in size in comparison

Table 7
Training effect heterogeneity by educational level.
Source: Microcensus 2005–2012, own calculations.

Local regressions	(1)	(2)
Panel A: Non-college		
Without covariates		
Treatment variable	0.0085	0.0158
	(0.0095)	(0.0133)
With covariates		
Treatment variable	0.0086	0.0160
	(0.0095)	(0.0133)
Panel B: College		
Without covariates		
Treatment variable	0.0844**	0.1383***
	(0.0354)	(0.0392)
With covariates		
Treatment variable	0.0831**	0.1377***
	(0.0353)	(0.0395)
Age FE	✓	✓
Year FE	✓	✓
Running variable	Linear	Quadratic

Notes: Standard Errors in parentheses; clustered at birth month level. Sample includes twelve month before and after reform cutoff. Significance levels: * p 0.10, ** p 0.05, *** p 0.01; Participation rates in training for cohort 1951 are 0.154, 0.317, 0.117 for all employed women, employed women with college and employed women without college, respectively. Number of observations: Panel A 22,984; Panel B: 5,535. Covariates: high initial level of education indicator (ISCED 97 Level 5-6), and a variable for medium education (ISCED 97 Level 3-4) as well as an indicator variable for west Germany. Source: Microcensus 2005-2012, own calculations.

et al., 2014). Other criteria and polynomials render bandwidths estimates ranging from 6 to 12. In all cases, our estimates remain in the same ballpark.

As additional robustness check, we examine if observations close to the cut-off drive our effects by estimating effects from donut-RD regressions. We estimate different specification of Eq. (2) without the one or two birth month closest to the cut-off on both sides. Online Appendix Table A.7 shows the resulting estimates for the linear and quadratic local polynomials as well as with and without individual control variables. In one specification (Panel A, column 2) the estimate is now only statistically significant at the ten percent level. All others remain highly significant, including from the more demanding two-month donut. Overall, this additional analysis confirms the main findings, with estimates of similar magnitudes throughout. Our results are not driven by outlier observations close to the cut-off.

6. Heterogeneity

6.1. Heterogeneity by initial education

We now extend the empirical analysis by focusing on effect heterogeneity along prior educational levels. We already showed descriptively that training participation positively correlates with prior educational levels in Section 2.2 and Figure A.1. In this analysis, we test if the reform effect also varies by prior educational level. Since we only focus on women close to the cut-off, to alleviate issues of sample size, we again aggregate the ISCED educational levels into two education groups of women classified as "college" and "non-college".

Before turning to the results, note that Geyer et al. (2018) show that employment effects of the same pension reform for women aged 60–62 are of similar size for highly educated women (9.5%) and for women without higher education (8.2%). Thus, any differences in the reform effect along the education dimension are not related to differences in retirement decisions.

For training before the age of 60 we find very strong differences by education for the different specifications (see Table 7). Looking at our preferred specification including a quadratic specification of the running variable, women with college education or more increase training

by nearly 13.8 percentage points, corresponding to a relative increase of about 43.4%.[31] The effect for women without college education is estimated to be close to zero and not significant at conventional levels.[32]

We tested (and rejected) in the main results section above the hypothesis that forward-looking employment effects mechanically create the training effects that we observe due to the sampling of our data. The finding that the positive training estimates are entirely driven by large and statistically significant effects for the college-educated now allows one additional test. If composition effects were driving the training effects, we would expect to find forward-looking employment effects for the same groups. In Online Appendix Table A.8 we present results for heterogeneity of employment effects. While the estimates are positive throughout, none of the estimates for college-educated workers is statistically significant. In contrast, two out of four estimates for workers without college are statistically significant at the 10% level or higher. This confirms our previous findings and interpretation.

6.2. Heterogeneity by firm size

Next, we examine if the training effect differs by the size of the company. For this, we estimate the RDD separately for women working in large vs small/medium-sized companies. We use the classification from the Microcensus, where companies with 50 or more employees are classified as large. Before turning to the interpretation of the results, we note again that company size is potentially an outcome in its own right, and can thus be considered endogenous. But as documented in Table 4, the pension reform did not have an affect on sorting into bigger companies.

Online Appendix Table A.9 splits the sample by company size. For large companies, looking at the preferred local regression approach, the effects vary between 0.79 and 2.65 percentage points when including controls. For small-medium sized companies the effects fall in the range of 3.22 and 4.82 percentage points. Overall, out of the eight estimates provided, five are significant. The estimates thus suggest slightly larger and more significant effects in small and medium sized companies, but the confidence intervals of the estimates are overlapping. We therefore conclude that the main dimension of heterogeneity is along levels of initial education.

7. Quantifying treatment effects of the reform

As discussed in Section 2.1, the pension reform only affected the working life of women who fulfill the eligibility criteria for the so-called *pension for women*. The Microcensus is a cross sectional data set without information about the employment history. Therefore, we cannot directly determine the eligibility within this data. As a result, our estimates should be interpreted as intention-to-treat (ITT) effects, giving a lower bound of the true effect.

To gauge information about actual eligibility, we use information from the SOEP longitudinal data, according to which about 76% of all women employed before entering retirement were indeed eligible for this pathway into retirement. This rate increases to approximately 89% when excluding self-employed and civil servants, who are not, by definition, eligible. Further, SOEP data show that about 86% of employed women without a college degree and 94% of women with a college degree fulfill the eligibility criteria.

With this information and the estimated effects (ITT) presented in Tables 2 and 7, we can derive the average treatment effect on the

---

[31] We discuss the magnitude of this effect in detail in Section 7.

[32] Note that Figure A.2 suggests that women aged 55 might be driving the strong results for women with a college degree. We therefore also estimate the heterogeneous results excluding age group 55 and the results remain unchanged.

Table 8
Average Treatment Effect on the Treated (ATT)

	(1)	(2)
Panel A: Whole sample		
ATT	0.0305	0.0443
Relative size ATT (in%)	19.81	28.77
Panel B: College		
ATT	0.0883	0.1463
Relative size ATT (in%)	27.85	46.15
Panel C: Non-college		
ATT	0.0099	0.0185
Relative size ATT (in%)	8.48	15.79
Running variable	Linear	Quadratic

Notes: The ATT is derived by weighting the ITT effects presented in Tables 2 and 7. The eligibility rates are calculated from the SOEP data. Participation rates in training for cohort 1951 are 0.154, 0.317, 0.117 for all employed women, employed women with college and employed women without college, respectively. The corresponding eligibility rates are 89.14, 94.12, 86.58. Number of observation Panel A: 28,519, Panel: B 22,984, Panel C: 5,535.

treated (ATT) overall and for the different educational groups. More precisely, we divide the ITT estimates by the respective eligibility share, E_g, for each educational group g to obtain ATT estimates:

$$ATT_g = \frac{ITT_g}{E_g}$$

These are presented in Table 8. Overall, the pattern of the ATT effects is similar to the ITT effects, but the effects are slightly larger. The point estimates suggest that overall probability to participate in training increases by roughly 4.4 percentage points in our preferred quadratic specification including covariates (Panale A, column 2). For women with a college degree the increase is substantially higher with 14.6 percentage points (Panel B, column 2), and the effect for women without college is close to zero (Panel C, column 2). These estimates imply a relative increase in the probability of training of 28.8% of all women and 46.2% for women with college degree.

These effects are sizeable so it is important to consider if they are plausible. We believe so, for three reasons. First, the effect sizes that we document are not dissimilar to the findings of the small existing empirical literature on training effects. In particular, Brunello and Comi (2015) estimate for private sector Italian workers effects on training of nine percent as a result of a policy that shifts the ERA by one year. In comparison, our estimates are larger but they are also estimated from a larger policy shift. Second, although our results represent sizable estimates, so far this discussion does not account for the fact that, while being distinct from zero at conventional levels of statistical significance, these coefficients are of course estimated with uncertainty. When we instead focus on the lower bounds of the corresponding 95%-confidence intervals, we obtain a value of 2.22 percentage points for the ITT using the local quadratic specification for the whole sample. This corresponds to an ATT of approximately 2.49 percentage points and a relative size of 16.17% overall. For college graduates the lower bound ITT estimate is 6.02 percentage points resulting in an ATT estimate of 6.40 percentage points, and a relative size of 20.18%, again using the quadratic specification. Last but not least, even lower hypothesized estimates cannot be refuted at conventional levels of statistical significance when using the results of the difference-in-discontinuities approach for the baseline.

8. Conclusion

In this paper, we provide causal evidence for the theory of human capital accumulation, which has the key prediction that education investments depend on the length of the payout period. We exploit a sizable pension reform that sharply increased the Early Retirement Age for women between two adjacent cohorts from 60 to 63 years. The analysis is based on the German Microcensus using RDD estimation approaches.

The empirical analysis offers support for the key prediction of human capital theory that the duration of the payoff matters for educational investment decisions. In more detail, our empirical results show that the increase in the ERA has a sizable effect on the human capital accumulation of employed women: in our preferred specification, the probability of training increases by about 4.4 percentage points, which corresponds to an increase of 28.8% for these age groups, and at least 16.2% considering lower bound of the 95% confidence interval. This finding is robust to changes in the bandwidth and for different specifications of the running variable in the RDD and is supported by placebo tests. Investigating heterogeneity, we show that the pension reform increases training for women with a college degree or more by more than fourteen percentage points, which corresponds to a relative increase of more than 46.2%, with a lower bound of 20.2%. The effect for women without college degree is not significant.

Our findings have important implications for the policy debate on pension reform and help to understand why an increase in the retirement age has positive employment effects but also leads to an increase in unemployment and inactivity (see, e.g. Geyer and Welteke, 2021). The debate usually abstracts from the dynamic human capital investment that we document. Our results show that retirement policy need not be thought of as independent of training decisions. Policy makers should consider positive effects on training demand that result from expansions of working life. Moreover, it remains an open question of why the training response is limited to highly-educated women. One explanation is that lowly-educated workers find it harder to increase their working life, thus benefiting less from potential training. A better understanding of the underlying reasons for the documented heterogeneity is key to increase training further, as well for the low educated to reduce inequality of old age income and the risk of old age poverty. From a policy perspective, in ageing societies increasing productivity or just the labor market attachment of older workers is becoming more critical. We believe, future work should examine the role of individual workers and firms in initiating the positive training effects that we document, adding to the still relatively underdeveloped literature on educational investments beyond initial schooling.

Data availability

The data that has been used is confidential.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.labeco.2023.102426.

References

Atalay, K., Barrett, G.F., 2015. The impact of age pension eligibility age on retirement and program dependence: Evidence from an Australian experiment. Rev. Econ. Stat. 97 (1), 71–87.

Baranov, V., Kohler, H.-P., 2018. The impact of AIDS treatment on savings and human capital investment in malawi. Am. Econ. J.: Appl. Econ. 10 (1), 266–306.

Barrett, A., O'Connell, P.J., 2001. Does training generally work? The returns to in-company training. ILR Rev. 54 (3), 647–662.

Bauer, A.B., Eichenberger, R., 2017. Endogenous aging: How statutory retirement age drives human and social capital. CREMA Working Paper Series 2017-02, Center for Research in Economics, Management and the Arts.

Becker, G., 1962. Investment in human capital: A theoretical analysis. J. Polit. Econ. 70.

Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. J. Polit. Econ. 75 (4, Part 1), 352–365.

Bertoni, M., Brunello, G., Mazzarella, G., 2018. Does postponing minimum retirement age improve healthy behaviors before retirement? Evidence from middle-aged Italian workers. J. Health Econ. 58, 215–227.

Bloom, D., Kuhn, M., Prettner, K., 2019. Health and growth. In: Oxford Research Encyclopedia on Economics and Finance.

Blundell, R., Dias, M.C., Goll, D.A., Meghir, C., 2019. Wages, experience and training of women over the lifecycle. NBER Working Paper, 25776.

Boersch-Supan, A., Wilke, C., 2004. The German public pension system: How it was, how it will be. NBER Working Paper 10525.

Brunello, G., Comi, S., 2015. The side effect of pension reforms on the training of older workers. evidence from Italy. J. Econ. Ageing 6, 113–122.

Calonico, S., Cattaneo, M.D., Farrell, M.H., Titiunik, R., 2018a. RDROBUST: Stata module to provide robust data-driven inference in the regression-discontinuity design. Statistical Software Components, Boston College Department of Economics.

Calonico, S., Cattaneo, M.D., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica 82 (6), 2295–2326.

Calonico, S., Cattaneo, M.D., Titiunik, R., 2018b. Manipulation testing based on density discontinuity. Stata J. 18 (1), 234–261.

Carta, F., de Philippis, M., 2021. Working horizon and labour supply: The effect of raising the full retirement age on middle-aged individuals. Bank of Italy Temi di Discussione (Working Paper) No. 1314.

Cattaneo, M.D., Titiunik, R., 2022. Regression discontinuity designs. Annu. Rev. Econ. 14 (1), 821–851.

Cervellati, M., Sunde, U., 2013. Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited. Econometrica 81 (5), 2055–2086.

Duggan, M., Singleton, P., Song, J., 2007. Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls. J. Public Econ. 91 (7), 1327–1350.

Fan, X., Seshadri, A., Taber, C., 2017. Understanding Earnings, Labor Supply, and Retirement Decisions. Michigan Retirement Research Center Working Paper 367.

Fitzpatrick, M.D., Moore, T.J., 2018. The mortality effects of retirement: Evidencefrom social security eligibility at age 62. J. Public Econ. 157, 121 – 137..

Foster, A.D., Rosenzweig, M.R., 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. J. Polit. Econ. 103 (6), 1176–1209.

Frazis, H., Loewenstein, M.A., 2005. Reexamining the returns to training:Functional form, magnitude, and interpretation. J. Human Resour. 40, 453–476.

French, E., 2005. The effects of health, wealth, and wages on labour supply and retirement behaviour. Rev. Econom. Stud. 72 (2), 395–427.

French, E., Jones, J.B., 2011. The effects of health insurance and self-insurance on retirement behavior. Econometrica 79 (3), 693–732.

Geyer, J., Haan, P., Hammerschmid, A., Peters, M., 2018. Labor market and distributional effects of an increase in the retirement age. IZA Discussion Paper 11618.

Geyer, J., Welteke, C., 2021. Closing routes to retirement: How do people respond? J. Human Recour. 56, 277–308.

Gohl, N., Kurz, E., Haan, P., Weinhardt, F., 2020. Working life and human capital investment. Iza discussion paper no. 12891, Institute of Labor Economics.

Gustman, A.L., Steinmeier, T.L., 1986. A structural retirement model. Econometrica 54 (3), 555–584.

Haan, P., Prowse, V., 2014. Longevity, life-cycle behavior and pension reform . J. Econometrics 178, Part 3, 582–601.

Jayachandran, S., Lleras-Muney, A., 2009. Life expectancy and human capital investments: Evidence from maternal mortality declines. Q. J. Econ. 124 (1), 349–397.

Killingsworth, M.R., 1982. 'Learning by doing' and 'investment in training': A synthesis of two 'rival' models of the life cycle. Rev. Econom. Stud. 49 (2), 263–271.

Kolesar, M., Rothe, C., 2018. Inference in regression discontinuity designs with a discrete running variable. Amer. Econ. Rev. 108 (8), 2277–2304.

Köller, O., Hasselhorn, M., Hesse, F., Maaz, K., Schrader, J., Solga, H., Spiess, C., 2017. Das bildungswesen in deutschland. Bestand und potenziale.

Kuhn, A., Wuellrich, J.-P., Zweimueller, J., 2010. Fatal attraction? Access to early retirement and mortality. IZA Discussion Papers 5160.

Leuven, E., 2005. The economics of private sector training: A survey of the literature. J. Econ. Surv. 19 (1), 91–111.

Manoli, D.S., Weber, A., 2016. The effects of the early retirement age on retirement decisions. In: Working Paper Series 22561, National Bureau of Economic Research.

Mastrobuoni, G., 2009. Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities. J. Public Econ. 93 (11), 1224–1233.

Montizaan, R., Coervers, 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.

Oster, E., Shoulson, I., Dorsey, E.R., 2013. Limited life expectancy, human capital and health investments. Amer. Econ. Rev. 103 (5), 1977–2002.

Picchio, M., van Ours, J.C., 2011. Retaining through training; even for older workers. Technical report, SSRN Electron. J..

Pischke, J.-S., 2001. Continuous training in Germany. Technical Report 3, Journal of Population Economics 523–548.

RDC of the Federal Statistical Office and Statistical Offices of the Laender, 2015. Microcensus, survey years 1996–2015, own calculations.

Ruhose, J., Thomsen, S.L., Weilage, I., 2019. The benefits of adult learning: Work-related training, social capital, and earnings. Econ. Educ. Rev. 72, 166–186.

Rust, J., Phelan, C., 1997. How social security and medicare affect retirement behavior in a world of incomplete markets. Econometrica 65 (4), 781–832.

Seibold, A., 2021. Reference points for retirement behavior: Evidence from german pension discontinuities. Amer. Econ. Rev. 111 (4), 1126–1165.

Soares, R.R., 2005. Mortality reductions, educational attainment, and fertility choice. Amer. Econ. Rev. 95 (3), 580–601.

Staubli, S., Zweimueller, J., 2013. Does raising the early retirement age increase employment of older workers? J. Public Econ. 108 (C), 17–32.

Zweimueller, J., Winter-Ebmer, R., 2000. Firm-specific training: Consequences for job mobility. IZA Discussion Papers 138, Institute for the Study of Labor (IZA).