Late-Career Workload Reduction and the Retention of Experienced Workers

Elena Castellaro*

November 19, 2026
Click here for the latest version

This paper examines whether easing workload at later career stages can retain experienced workers in demanding occupations facing persistent shortages. I study a Dutch program that subsidizes reduced teaching hours for older secondary-school teachers, exploiting a 2014 reform that restricted eligibility by date of birth in a regression discontinuity design. Using administrative records on the universe of Dutch teachers, I find that eligibility raises take-up but has no detectable effect on retention. This null masks different responses by pre-reform contract size, a proxy for whether teachers had already reduced their hours at their own expense. Full-time teachers are 6.1 percentage points more likely to continue teaching, and the additional hours supplied through increased retention largely offset the hours they give up through the program. Part-time teachers take up the subsidized reduction without staying longer, so their teaching supplied falls. Eligibility also appears to improve physical health, with linked health-insurance claims showing a decrease in pain-medication prescriptions. The results suggest that subsidized workload reduction retains mainly workers for whom it relaxes a binding constraint on hours.

Keywords: workload reduction, retention, older workers, hours constraints, teacher shortages, health
JEL codes: J22, J26, J45, J63, J81, I21

*University of Amsterdam and Tinbergen Institute. Email: e.castellaro@uva.nl.
I would like to thank my supervisors, Monique de Haan and Bas van der Klaauw, for their advice and support. I am also grateful to Paola Bertoli, Zichen Deng, Didier Fouarge, Julien Grenet, Julius Ilciukas, Ingo Isphording, Maarten Lindeboom, David Margolis, Eric Maurin, Hessel Oosterbeek, Erik Plug, Sara Signorelli, Maddalena Totarelli, Mark Witte, Esmée Zwiers, as well as the participants at various conferences and seminars, for their helpful comments and feedback. The non-public microdata used in this paper are available via remote access to the Microdata services of Statistics Netherlands (project agreement 10348).

1 Introduction

Workers in physically or mentally demanding occupations often leave their jobs well before retirement age. Prolonged work-related strain can accelerate health decline (Ravesteijn et al., 2018) and drive workers out of the labor force or into less demanding jobs (Filer and Petri, 1988; Blundell et al., 2023). These exits are particularly costly in demanding public-service occupations such as teaching, nursing, and elder care, where persistent shortages make experienced workers difficult to replace (OECD, 2023). Retention policies in these occupations often rely on financial incentives (Clotfelter et al., 2008; Falch, 2011; Ruffini, 2024), yet if work-related strain contributes to early exits, improving working conditions may offer an alternative route to retention. This paper examines whether reducing workload at later career stages helps retain experienced workers in demanding professions.

I study this question in the context of Dutch secondary education, exploiting a national program (BAPO, "Promotion of Labor Participation of Older Workers") that allowed teachers aged 52 and older to reduce their annual teaching load by up to 10 percent through subsidized leave. The program covered roughly 65 percent of the forgone salary, while teachers retained full pension rights. Two features make teaching a natural setting for this question. First, work-related strain is particularly acute in the profession (OECD, 2025). Dutch teachers aged 50–65 report burnout complaints nearly twice as often as other workers (30 versus 17 percent, Netherlands Working Conditions Survey, NEA), and workload is one of the main reasons they give for leaving (Van Casteren et al., 2023). Second, exits are difficult to absorb: in the Netherlands, outflows persistently exceed inflows, and unfilled positions are often covered by substitutes not qualified in the subject (Adriaens and den Uijl, 2025; Herrmann and Rockoff, 2012).

A subsidized reduction in teaching hours may retain teachers who would otherwise leave the profession because of high workload. Although Dutch teachers can reduce their contracted hours without a subsidy, doing so is costly: they forgo wages and pension accrual, may be unable to regain those hours later, and may have their request denied. The subsidy should therefore matter most for teachers who would prefer fewer hours but have not reduced them because of these costs, and less for teachers who already work the hours they prefer. Contract size provides a proxy for this distinction: part-time teachers are more likely to have already adjusted their hours to their preferences, whereas full-time teachers may include workers who would prefer fewer hours but have not made that adjustment. Consistent with this interpretation, late-career full-time teachers report greater work pressure than their part-time counterparts and are more likely to say

they would work to a higher age if they could work fewer hours (Netherlands Working Conditions Survey, NEA: Table C.1).

Estimating the causal effect of subsidized workload reduction on retention is complicated by voluntary program participation. Participants may differ from non-participants in job strain, health, and attachment to the profession, which can independently affect retention. To address this selection problem, I exploit a 2014 reform that raised the eligibility age for subsidized workload reductions from 52 to 57, with the new age applying only to teachers born after a cutoff date. Teachers born just before the cutoff (the treatment group) remained eligible from age 52, whereas those born just after it (the control group) became eligible only at 57. I use a regression discontinuity (RD) design around the birth-date cutoff to estimate the causal effect of access to the program from age 52 rather than 57 (henceforth, eligibility), over the years before the control group becomes eligible.

I use administrative data on the universe of Dutch secondary-school teachers, linked through individual identifiers to demographic characteristics, labor-market outcomes, and mandatory health-insurance claims. The claims data record prescriptions for antidepressants and pain medications, which I use as proxies for mental and physical health, as well as total healthcare expenditure. Since health-insurance claims cover the entire resident population, health outcomes are observable even when teachers leave the profession.

I find that eligibility increases BAPO take-up by 40 percentage points but has no detectable effect on retention over 2015–2018. This aggregate null, however, masks substantial heterogeneity by contract size. While take-up rises among both full-time and part-time teachers, retention increases only among full-time teachers. For part-time teachers, the estimated effect on retention is negative but statistically insignificant. Among full-time teachers, eligibility raises the probability of remaining in teaching by 6.1 percentage points, preventing approximately 60 percent of the exits that would otherwise have occurred. The effect emerges in the first year of eligibility and persists through 2024, even after the control cohort reaches the eligibility age of 57. The destination estimates are imprecise, but suggest that more than half of the retained teachers would otherwise have moved to another job, mostly outside education, while the remainder would have left employment.

The net effect of the program on total teaching supplied depends on whether the hours gained through higher retention offset the hours withdrawn through leave. In the full sample, the estimated effect over 2016–2018 is small and statistically insignificant, reflecting two opposing responses. Among full-time teachers, retention gains largely offset

the leave: the estimate is positive, and its confidence interval rules out more than a small decline. Part-time teachers take up the leave without remaining in teaching longer, so their teaching supplied falls by about 10 percent relative to the control mean.

Reducing teaching hours also appears to improve physical health. Eligibility reduces pain-medication use by 5.7 percentage points, consistent with a decline in musculoskeletal problems, which are a common occupational complaint among teachers (Erick and Smith, 2011). I find no detectable effect on antidepressant use or healthcare expenditure. The decline in pain-medications use is statistically significant only among part-time teachers, whose retention does not change, so for this group the estimate largely isolates the effect of reduced hours. Among full-time teachers, the additional years in teaching entail no detectable health cost, although the estimates are too imprecise to rule out moderate effects.

The contract-size split suggests that the subsidy raises retention only among teachers who would prefer fewer hours but have not reduced them because of the costs. It does not, however, identify the channel through which a cheaper reduction keeps them in teaching. Fewer hours may relieve strain and make a demanding job sustainable, or teachers may value the leave as additional leisure. The two channels yield different predictions: if leave is valued for its own sake, the retention response should not depend on strain. I test this by applying the same design to full-time support staff, who are entitled to the same program but report substantially lower work pressure, and comparing them with full-time teachers in the lowest salary scale, whose earnings are closest to those of support staff. Eligibility raises take-up by more than 30 percentage points in both groups, but retention increases only among teachers. The evidence thus suggests that the program raises retention by relieving strain rather than by subsidizing leisure.

When an eligible teacher stays but takes leave, her school must cover a fraction of her timetable rather than a full vacancy. Exploiting cross-school variation in exposure to eligible cohorts, I find that the hours seniors give up are covered by younger incumbents expanding their contracts, with no detectable change in hiring.

The main benefit of retaining teachers is avoiding vacancies, which Dutch secondary schools often fill with out-of-field teachers. While the program does not pay for itself financially, accounting for the resulting student achievement gains using estimates from the teacher value-added literature yields a positive net return. This return is driven entirely by full-time teachers, suggesting that restricting eligibility to this group would substantially reduce program costs while preserving its benefits.

Related literature This paper contributes to three literatures. The first evaluates policies to retain workers in public-service occupations facing shortages. Among teachers, higher pay lowers turnover (Clotfelter et al., 2008; Falch, 2011; Hendricks, 2014) and shapes sorting across districts (Biasi, 2021); pension incentives govern late-career exit (Brown, 2013); and targeted bonuses raise early-career retention (Sims and Benhenda, 2022). Higher wages similarly reduce staff turnover in nursing homes (Ruffini, 2024). Causal evidence on working conditions is scarce. The two existing studies of workload reductions for older teachers, a free 5.8 percent cut in teaching load at age 55 in Norway (Bratberg et al., 2020) and the pre-reform BAPO scheme in the Netherlands (Leuven, 2006), focus on sickness absence and are identified from an age threshold. I instead estimate effects on retention, comparing teachers of nearly the same age who differ only in access to the reduction, and show that the reduction retains teachers by relieving strain rather than by subsidizing leisure.

The second concerns the cost of adjusting hours within a job. Hours are often constrained at the current employer (Lachowska et al., 2026), and workers who want to work fewer hours tend to change jobs rather than contracts (Altonji and Paxson, 1988, 1992; Euwals, 2001; Fouarge and Baaijens, 2009). For older workers, the alternative to a reduction is often exit from employment (Gielen, 2009; Ameriks et al., 2020). Subsidized hours reductions aim to lower this cost, but most evaluations find that participants reduce hours without delaying retirement, leaving total labor supply unchanged or lower (Graf et al., 2011; Albanese et al., 2020; Börsch-Supan et al., 2018; Bensnes et al., 2026), while Rutten et al. (2025) find an increase among Dutch civil servants. Losses arise when participants would have kept working anyway (Graf et al., 2011). I show that gains arise where hours are constrained: the subsidy retains full-time teachers, offsetting the leave they take, while part-time teachers, who had already reduced their hours, take up the subsidy without staying longer. I also study an earlier margin, at which the alternative to a demanding job is often another job rather than retirement.

Finally, I contribute to the literature on working hours and health. Evidence that reducing hours improves health comes mainly from statutory workweek reductions covering the whole workforce (Ahn, 2016; Sánchez, 2017; Cygan-Rehm and Wunder, 2018; Berniell and Bietenbeck, 2020), which find at most modest effects on self-reported health. Because these reforms cut contractual hours without changing the work itself, they cannot separate fewer hours from a lighter workload. In the Norwegian teacher reform, Bratberg et al. (2020) find lower sickness absence and better mental health, concentrated among men in poor prior health, but no effect on musculoskeletal problems. I instead find a

reduction in pain-medication use, consistent with improved musculoskeletal health.

The remainder of the paper proceeds as follows. Section 2 describes the institutional setting and the 2014 reform. Section 3 presents the data and Section 4 the empirical strategy and its validation. Section 5 reports the effects of eligibility on take-up, retention, teaching supplied, and health, together with robustness checks. Section 6 examines whether the retention response reflects strain relief or the value of leave as leisure. Section 7 studies how schools cover the hours released through leave and assesses the program's cost-effectiveness. Section 8 concludes.

2 Institutional Context

2.1 Dutch Secondary Education

Dutch secondary education (*voortgezet onderwijs*, VO) consists of three tracks that differ in length and in the tertiary education they give access to: VMBO (four years, prevocational), HAVO (five years, leading to universities of applied sciences), and VWO (six years, leading to research universities). Schools recruit and employ staff autonomously, but wages, pay scales, and working-time arrangements are set sector-wide through the collective labor agreement (CAO-VO), negotiated between school boards and unions. Personnel is formally divided into management, teaching staff, and support staff, a legal distinction grounded in qualification requirements. Teaching staff hold a statutory teaching qualification and bear formal responsibility for the classroom, whereas support staff face no qualification requirement and sit on separate salary scales.[1] In addition to directly employed teaching staff, schools hire teachers through temporary agencies or as self-employed contractors. Schools use this channel mainly to cover sickness absence and unplanned vacancies, and it accounted for 2.3 percent of personnel costs in 2013 and 3.7 percent in 2015 (Leemans et al., 2017). For teaching staff, a full-time appointment corresponds to a 40-hour week, of which 22 hours are spent teaching in 50-minute periods; the remainder covers preparation, grading, and meetings. Teachers retain considerable control over their working time, as they may request annual adjustments to their contractual full-time equivalent (FTE). Part-time work is therefore widespread: the average teacher held a 0.83 FTE contract in 2013 (Table 1). Despite this flexibility, the sector appears to be under considerable strain. Data from the Netherlands Working Conditions Survey (NEA; see Appendix C) show that, compared with the Dutch workforce as a whole, VO

[1] Support staff include teaching assistants, technical and ICT staff, and administration. Teaching assistants may work directly with classes, but only under the responsibility of a qualified teacher.

teachers are more likely to report high workload, high emotional demands, and burnout complaints, and more likely to want to reduce their working hours in order to keep working for more years (Figure A.1). These differences are present in every age group between 50 and 65.

2.2 BAPO and the 2014 Reform

Since 1994, teachers aged 52 and older had been eligible for BAPO (*Bevordering Arbeidsparticipatie Ouderen*, "Promotion of Labor Participation of Older Workers"), a subsidized leave program that allowed them to reduce their teaching hours at a reduced cost. Schools subsidized 65% of the salary associated with the leave, while pension accrual continued on the full contract. Teachers could also reduce their hours without BAPO, but this lowered both salary and pension accrual and could make hours difficult to regain.[2]

Schools financed the subsidy from their lump-sum grant, which varies with enrollment but not with the age of staff.[3] As the workforce aged, program costs increased and the scheme came under political pressure.[4] The 2014 reform of the CAO-VO responded by raising the minimum age for subsidized leave from 52 to 57. The reform was announced in April 2014 and took effect on 1 August of that year. It included a transitional arrangement, which provides the treatment studied here: teachers born before 1 August 1962, who had turned 52 by the 31 July 2014 reference date and had been continuously employed by a publicly funded educational or research institution since 1 August 2009, retained access to subsidized leave from age 52. Teachers born on or after 1 August 1962 became eligible only at age 57, a five-year deferral.[5] Figure 1 plots the eligibility profiles of the two cohorts (blue and orange lines). I exploit this discontinuity in a regression discontinuity design, comparing teachers born just before (treated) and just after the cutoff (control) to estimate the causal effect of access to the program during the five-year window.

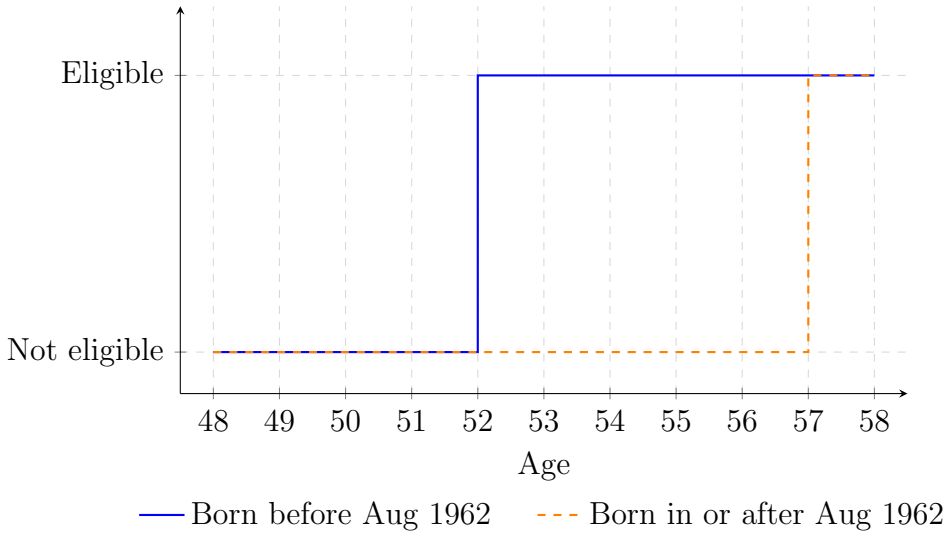
[2]Teachers may request a contract adjustment once a year (VO-raad, 2014, art. 6.1), but increases may be refused due to a lack of work, formation space, or personnel budget (*Wet aanpassing arbeidsduur*, art. 2).

[3]The grant's personnel component multiplies a pupil-driven number of normative staff positions by a flat national wage norm per school type and contains no BAPO component (Algemene Rekenkamer, 2014).

[4]A 2011 parliamentary motion urged that the scheme be phased out (Tweede Kamer, 2011), and an assessment of the teacher labor market found high take-up but little reduction in sickness absence (Leuven, 2006).

[5]From age 57, teachers were allowed to obtain an additional 170 hours of leave under certain circumstances.

Figure 1: Eligibility for the BAPO leave program after 1 August 2014



Notes: Eligibility for subsidized BAPO leave by age under the 2014 transitional arrangement. Teachers born before 1 August 1962 remain eligible from age 52; those born on or after that date become eligible only at 57.

Under the transitional arrangement, a full-time teacher in the treatment group was entitled to 50 hours of leave without co-payment and 120 hours with a 50 percent co-payment, for 170 hours in total.[6] This amounts to about 10 percent of the 1,659 annual working hours of a full-time contract and corresponds to a reduction of three 50-minute lessons per week (VO-raad, 2014).[7] The leave entitlement scales with contract size, as do all rights under the CAO-VO (VO-raad, 2014), while pension accrual and other salary-linked entitlements are based on the unreduced salary. BAPO leave cannot fall below the previous year's amount at a constant contract size, and to ensure that the leave is not taxed as early retirement, a teacher must continue working at least half of their pre-leave weekly hours (VO-raad, 2014).[8]

[6] For support staff, the co-payment was 40 percent.

[7] For a full-time teacher taking all 170 hours, the co-payment is $0.5 \times 120/1{,}659 \approx 3.6$ percent of annual salary. The price of leave was essentially unchanged: under the old scheme, with a 35 percent co-payment on all 170 hours, leave hours were paid at 65 percent of salary; under the transitional arrangement, the corresponding rate was $(50 + 0.5 \times 120)/170 \approx 0.65$.

[8] BAPO alone never breaches this floor, which binds only in combination with other leave. Since 2005, an employer has owed a final levy (52 percent from 2011) on payments under schemes that bridge income to the pension date; reduced-hours arrangements for older workers escape it only if the employee continues working at least half of the original contract (*Wet op de loonbelasting 1964*, art. 32ba; *Uitvoeringsregeling loonbelasting 2011*, art. 8.7).

Teachers in the control group had no access to subsidized leave before 57. All they received after the 2014 reform was the 50-hour *basisbudget* granted to every employee, which could be used for task reallocation, training, or leave (VO-raad, 2014).[9] The difference between the cohorts at age 52 is thus the 120 co-paid hours, scaled to contract size. The transitional arrangement and *basisbudget* differ in usability, however. BAPO hours are taken as leave, whereas use of the *basisbudget* depends on school-level task allocation: among staff who do not use it for training or task reallocation, 52 percent report that their school offers no opportunity to take the hours, while school leaders cite lesson coverage and limited replacement capacity as barriers (Voion, 2026). I therefore interpret the estimates as the effect of the BAPO package as a whole rather than of 120 hours of leave alone. For consistency, I use "BAPO" to refer to the entitlement under both the pre-reform rules and the transitional arrangement.

3 Data and Sample Selection

I use Dutch administrative data from Statistics Netherlands, which link secondary-school teachers to demographic characteristics, labor-market outcomes, and health records through unique individual identifiers. The teacher records come from the Education Executive Agency of the Ministry of Education (DUO) and cover the universe of personnel employed in publicly funded secondary schools between 2006 and 2024. Dutch secondary schools are predominantly publicly funded, with privately funded schools accounting for only a small share of the sector.[10] Although the registers also cover management and support staff, I restrict the main analysis to teaching staff. The registers record information on each employment relationship as observed in October of each year, including contract type, salary scale, gross monthly earnings, and annual full-time equivalents (FTE), and, from 2016 onward, the amount of leave taken under BAPO. Contract size defines the two groups analyzed throughout: following the Statistics Netherlands definition, I classify teachers as full-time if their 2013 contract is at least 0.95 FTE and as part-time otherwise. I complement the school-reported registers with payroll-based employer–employee records, which identify participation in any subsidized leave program involving a salary reduction and provide annual analogues of the October snapshot variables. Finally, I observe each teacher's gender, immigrant background, and exact day of birth.

[9] Pro rata to the contract and with pension accrual unchanged. Teachers in the treatment group could forgo BAPO leave and receive the *basisbudget* instead.

[10] Schools are legally required to report educational personnel to DUO annually. Coverage is near complete: between 2009 and 2020, an average of 2 percent of secondary-school teachers went unreported.

I further link teachers to claims data from the Dutch mandatory health insurance, which covers all residents. The claims include reimbursed prescription medications dispensed outside hospitals and total healthcare expenditure. Health is a natural outcome of the program, whose purpose is to make teaching sustainable at older ages by easing late-career workload. The strain it targets has both a psychological and a physical component: teachers frequently report psychological complaints (OECD, 2020; van den Heuvel and de Vroome, 2023), and musculoskeletal problems are among the most common physical complaints in the workforce (TNO, 2024). I therefore use antidepressant and pain-medication prescriptions as measures of mental and physical health, respectively. Because the claims cover all residents, health is observed regardless of subsequent employment, for teachers who leave teaching as well as for those who stay.

The main sample consists of teachers employed in secondary education in 2013 for whom teaching was the main source of income in that year, and born between December 1961 and March 1963.[11] Teachers born between December 1961 and July 1962 qualified for the transitional BAPO arrangement and form the treatment group; those born between August 1962 and March 1963 became eligible only at age 57 and form the control group. The resulting sample contains 2,278 teachers, whom I follow from 2015 to 2018 in a balanced panel of 9,112 person-year observations.

I impose two further restrictions. First, I drop 43 teachers who died or emigrated during the panel, whose health outcomes are unobserved; eligibility does not affect the probability of ceasing to be a Dutch resident (Table A.1), so the restriction is unlikely to induce differential selection at the cutoff. This restriction leaves a balanced panel of 2,235 teachers and 8,940 person-years. Second, 2.2 percent of person-years appear in the employer–employee data under an education-sector employer but are absent from the teacher register in the same year, typically because of delayed or erroneous school reporting that the October snapshot misses. Including these observations would require assuming continued employment as a teacher, which might not hold if the individual moved into a support role within the same school. I therefore exclude these observations from the main specification. The final estimation panel is therefore unbalanced, with 8,747 person-years for the same 2,235 teachers.[12] Table A.2 shows that the results are unchanged when these person-years are instead imputed as teaching employment and the

[11] This corresponds to a symmetric eight-month (240-day) bandwidth around the 1 August cutoff. I fix a common bandwidth across outcomes, set to the median of the outcome-specific MSE-optimal bandwidths of Calonico et al. (2014) computed with a uniform kernel, so that all estimates are identified from the same set of teachers.

[12] Appendix B describes the sample construction in detail.

panel is kept balanced.

Table 1: Descriptive statistics

	Whole sample		Full-time		Part-time	
	Mean	SD	Mean	SD	Mean	SD
Panel A: Demographics						
Female	0.55	0.50	0.31	0.46	0.73	0.45
Immigrant background	0.16	0.37	0.16	0.36	0.17	0.37
Age	50.9	0.47	50.9	0.47	50.9	0.48
Panel B: Labor market						
Contract size (FTE)	0.83	0.19	1.01	0.03	0.70	0.15
Monthly salary (euros)	3,456	983	4,273	541	2,850	777
Panel C: Health						
Antidepressants	0.06	0.24	0.04	0.21	0.08	0.27
Pain medication	0.18	0.39	0.18	0.38	0.18	0.39
Healthcare expenditure (euros)	1,715	5,938	1,496	5,265	1,877	6,387
Observations	2,235		952		1,283	

Notes: Analysis sample in 2013, the year before the reform: teachers employed in Dutch secondary education and born within 240 days of the 1 August 1962 cutoff. Labor-market variables come from the DUO teacher register (October); health variables from basic health-insurance claims (calendar year). Monetary values in 2024 euros. Full-time is ≥ 0.95 FTE in 2013.

3.1 Variable Construction

Labor-market outcomes come from two administrative sources. I measure *BAPO take-up* using employer–employee records, which report annual participation in any subsidized leave arrangement involving a salary reduction. Given the age composition of the sample, this measure almost exclusively captures BAPO.[13] The remaining outcomes come from the teacher register. *Retention* is an indicator equal to one if an individual is observed as teaching staff in a given year and zero if she is absent from the register or observed in a non-teaching occupation. Moves to temporary agency or self-employed work are classified as exits. From the perspective of school boards, these arrangements represent exits from direct employment: agencies are used primarily when vacancies cannot be filled directly,

[13]The measure may also include generic unpaid leave classified as "other" (*overig verlof*), which is distinct from maternity and sick leave.

at higher cost and typically for short spells. In practice, this channel is limited, accounting for 2.3 percent of personnel costs in secondary education in 2013 and 3.7 percent in 2015 (Leemans et al., 2017).

From 2016 onward, I also observe the FTE share allocated to BAPO leave and construct the *leave share*, defined as leave FTE divided by contractual FTE, for teachers who remain in the teacher register. This is the fraction of a teacher's contract taken as subsidized leave; under the collective agreement it cannot exceed 10 percent. Finally, *total teaching time* measures actual teaching hours in FTE, with teachers who have left the register coded as zero. This outcome captures the program's net effect on teaching supplied, combining reductions in hours through leave with changes in retention.

I measure health along several dimensions. *Antidepressants* is an indicator of mental health that equals one if an individual fills any reimbursed prescription for an antidepressant (ATC code N06A) in a given year and zero otherwise (Huber et al., 2013). Because antidepressants can be prescribed in primary care without a formal psychiatric diagnosis, this measure can capture psychological distress that may not appear in diagnosed mental healthcare. I therefore also construct a broader measure of mental health that equals one if an individual either fills an antidepressant prescription or has any reimbursed expenditure on generalist basic mental healthcare.[14]

Pain medication is an indicator of physical health that equals one if an individual fills any reimbursed prescription for analgesics or anti-inflammatory drugs in a given year (ATC codes N02A, N02B, or M01A).[15] Because most M01A drugs are available over the counter in the Netherlands, the measure captures pain that is sufficiently severe to generate a reimbursed prescription rather than everyday discomfort.[16]

Finally, I use annual total healthcare expenditure, deflated to 2024 prices, as a summary measure of healthcare use. Expenditure is highly skewed, and its mean is driven by hospital episodes largely unrelated to workload. Because nearly all teachers have some claims each year, an indicator for any expenditure would also carry little information.

[14] Dutch mental healthcare is organized in echelons. Mild complaints are treated within general practice by a mental health practice nurse, whose services are financed within the GP segment and are not separately identified in the claims data. Patients with more serious complaints are referred either to generalist basic care, for mild to moderate disorders, or to specialist care, for severe or complex disorders. I exclude specialist care because generalist basic care captures the margin of psychological distress that a reduction in workload could plausibly affect.

[15] Huber et al. (2013) use N02A and N02B only. I additionally include M01A (anti-inflammatory and antirheumatic products), which are commonly prescribed for neck and back pain associated with the physical demands of teaching. Results are robust to the narrower definition.

[16] This restriction implies that the measure does not capture pain treated exclusively with over-the-counter medication.

I therefore code an indicator for expenditure above the median of the 2013 distribution in the analysis sample, which separates years of routine care from years with additional treatment, the margin at which work-related illness is most likely to appear.

Table 1 presents descriptive statistics for the analysis sample in 2013, the year before the reform, for the full sample and separately for full- and part-time teachers. Teachers are 51 years old on average, and 16 percent have an immigrant background.[17] The two groups differ substantially in gender composition: women account for 31 percent of full-time teachers, compared with 73 percent of part-time teachers. Full-time teachers hold contracts of 1 FTE and earn average gross monthly salaries of 4,273 euros, compared with 0.70 FTE and 2,850 euros among part-time teachers. Antidepressant use is twice as high among part-time teachers (8 versus 4 percent), consistent with their higher share of women. Pain-medication use and annual healthcare expenditure are similar across the two groups.

4 Empirical Design

This paper studies the effects of eligibility for subsidized late-career workload reduction on teachers' outcomes. For identification, I exploit the 2014 BAPO reform in a regression discontinuity (RD) design that compares teachers born on either side of an eligibility cutoff. Teachers born before 1 August 1962 had turned 52 by the reform date and retained access to BAPO from that age, whereas those born on or after that date were still 51 and, under the post-reform rules, could take up the leave only from age 57. Date of birth thus generates a sharp discontinuity in program access.

I use the panel structure of the data to estimate the following baseline RD specification:

$$Y_{it} = \alpha + \beta D_i + \gamma(R_i - c) + \delta\, D_i(R_i - c) + X_i'\eta + \lambda_t + \varepsilon_{it}, \tag{1}$$

where Y_{it} is the outcome of teacher i in year t, R_i is teacher i's date of birth expressed in days, c is the cutoff (1 August 1962), and $D_i = \mathbf{1}\{R_i < c\}$ indicates eligibility for BAPO from age 52. The interaction $D_i(R_i - c)$ allows the slope to differ on each side of the threshold, X_i is a vector of baseline covariates and λ_t are year fixed effects. [18] Because I observe the same individuals over multiple years, standard errors are clustered at the individual level. The coefficient of interest, β, measures the discontinuity in outcomes

[17] An individual or at least one parent was born abroad.

[18] The covariates are measured in 2013 and include gender, immigrant background, FTE, and antidepressant and pain-medication use.

at the cutoff and identifies the effect of BAPO access at age 52. The main specification excludes covariates.

The analysis window spans 2015 to 2018. I exclude 2014 because the treatment group is only partially exposed in that year: eligible teachers could enter BAPO under the pre-reform rules upon turning 52, so those who entered had already used it for a few months by the time outcomes are measured in October 2014. The window ends in 2018, the final year before the control group turns 57 and becomes eligible under the post-reform rules. The main specification pools the four post-reform years to gain precision, and I also report year-by-year estimates to trace the dynamics of the response. To assess whether the effects persist once the control group also gains access, I additionally extend the window to 2024, when the treatment group reaches age 62. I estimate equation (1) by local linear regression within a 240-day bandwidth around the cutoff, using a uniform kernel. I hold the 240-day bandwidth fixed across outcomes to improve the comparability of the estimates and to limit concerns that bandwidth selection drives the results. The chosen window is in line with the median MSE-optimal bandwidths (Calonico et al., 2014) computed separately for the main labor-market and health outcomes. I assess sensitivity to this choice in Section 5.5.

Heterogeneity by Contract Size The program should not affect all teachers equally. Part-time work is common in Dutch secondary education: the average contract was 0.83 FTE in 2013, reflecting both teachers hired on part-time contracts and teachers who later reduced their hours.[19] Working fewer hours is costly in either case. The teacher forgoes the full wage and permanently lowers pension accrual on the hours given up, and hours once given up are difficult to regain. A later reduction may also be refused on scheduling or staffing grounds.[20] Moreover, contracts cluster at whole working days (37 percent of teachers are at 1.0 FTE and 15 percent at 0.8, with smaller clusters at 0.7, 0.6, and 0.5), so leaving full-time typically means giving up a day, twice the reduction BAPO provides. BAPO removes these costs. It is an annual entitlement that employers cannot refuse, taken within the existing contract, with salary-linked rights calculated on the unreduced salary (VO-raad, 2014). The design cannot separate which of these costs binds for a given teacher.

For a teacher who bears high workload and would prefer fewer hours but for whom a reduction is too costly or not granted, the choice may be between full hours and leaving

[19] Among teachers born in 1960, 33 percent reduced their contractual hours at least once between ages 46 and 60 (author's calculation from the DUO personnel register).

[20] *Wet aanpassing arbeidsduur*, art. 2.

teaching. Eligibility adds a third option, a reduction without these costs, and may keep her in teaching. A teacher already working her preferred hours at the unsubsidized price still benefits from the cheaper leave, but she would have stayed anyway, so the program changes how much she works rather than whether she stays. Contract size helps distinguish the two types. A part-time contract indicates a teacher who accepted the cost of fewer hours, at hiring or later, and is therefore more likely to work her preferred hours. A full-time contract pools teachers content with their hours and teachers who would prefer fewer but have not reduced them.

I therefore report all estimates separately for full-time and part-time teachers, classified by their 2013 contract, before the reform was announced, using the Statistics Netherlands threshold of ≥ 0.95 FTE.[21] The retention response should be concentrated among full-time teachers, whereas among part-time teachers the program should reduce hours without raising retention. Two further results support this interpretation. Ineligible full-time teachers reduce their contracts at their own expense, precisely the costly adjustment the program replaces (Appendix D). The full-time/part-time gap also persists within the lowest salary scale, where pay is roughly constant (Section 6).

Continuity assumption Identification rests on the standard RD continuity assumption (Cattaneo et al., 2020a): in the absence of the reform, potential outcomes would vary smoothly with date of birth at the cutoff. A discontinuity at 1 August 1962 can therefore be attributed to the reform. The assumption could fail if date of birth were systematically related to unobserved determinants of late-career labor-market or health outcomes, for example through season-of-birth effects on early-childhood health and human capital (Buckles and Hungerman, 2013).

I assess this assumption in two ways. First, I estimate equation (1) using pre-reform values of the main labor-market and health outcomes, separately for the full sample and by contract type. Because these outcomes are measured before teachers become eligible for BAPO, a discontinuity at the cutoff would indicate that birth date predicts outcomes for reasons unrelated to the reform.[22] I find no statistically significant discontinuity for any outcome (Table A.4), and the corresponding RD plots show no visible jump at the cutoff (Figure A.2). Second, I test for discontinuities in predetermined characteristics, including gender and immigrant background. These characteristics are balanced at the

[21] Results are robust to alternative thresholds (Table A.3).

[22] Teachers born in December 1961 turned 52 during 2013 and were in principle entitled to BAPO. Eligibility begins in the month following the 52nd birthday, however, so no teacher within the estimation bandwidth could take up the program in 2013.

cutoff for both full-time and part-time teachers (Table A.4; Figure A.3). The largest discrepancy is a 6 percentage point higher share of women to the left of the cutoff in the pooled sample, which is not statistically significant. Together, these results support the continuity assumption.

No sorting around the cutoff A valid RD design also requires that individuals cannot manipulate the cutoff. This concern is negligible in this setting because the running variable was fixed decades before the reform, and the cutoff was introduced retroactively in the 2014 collective agreement. Consistent with this, Figure A.4 shows a smooth density of the running variable with no evidence of bunching around the threshold. A manipulation test following Cattaneo et al. (2020b) yields no evidence of sorting at the cutoff ($p = 0.204$).

No anticipatory effects or confounding policies Although the running variable cannot be manipulated, the composition of the sample could still respond to the reform if teachers adjusted their employment in anticipation of losing access to BAPO. I therefore restrict the sample to teachers employed in secondary education in 2013, before the reform was announced, so that the sample is determined prior to any behavioral response.[23] Finally, the estimates would be confounded if another policy affecting my sample used 1 August 1962 as an eligibility cutoff. I am not aware of any such Dutch policy.

5 Results

This section reports estimates of equation (1) for the full sample and separately for full- and part-time teachers. I examine take-up, retention, teaching supplied, and health, over 2015–2018 (2016–2018 for teaching supplied), the years in which eligibility differs across groups, and then extend the window through 2024, when both groups are eligible. RD plots show mean outcomes in 30-day birth-cohort bins, with linear fits on either side of the cutoff.

[23] The reform was first set out in the CAO-VO 2014–2015 Negotiators' Agreement, concluded on 15 April 2014, which abolished the BAPO scheme and introduced a new seniority arrangement. While discussions about reform had taken place prior to this date, the exact rules (including the increase in the eligibility age from 52 to 57 and the transitional arrangement based on the 31 July 2014 reference date that effectively determined the cutoff) were only specified in the April agreement.

5.1 Take-up

Eligibility increases BAPO take-up by 40 percentage points in the full sample (Table 2; Figure 2), with larger effects for full-time than part-time teachers (47 versus 34 percentage points). Figure A.5 reports year-specific RD estimates for 2010–2020. The gap emerges in 2014 and narrows in 2019, when the control group becomes eligible at age 57, and it does not close fully, consistent with gradual take-up and with deferral reducing the pool of potential participants through exits or unsubsidized reductions in hours.

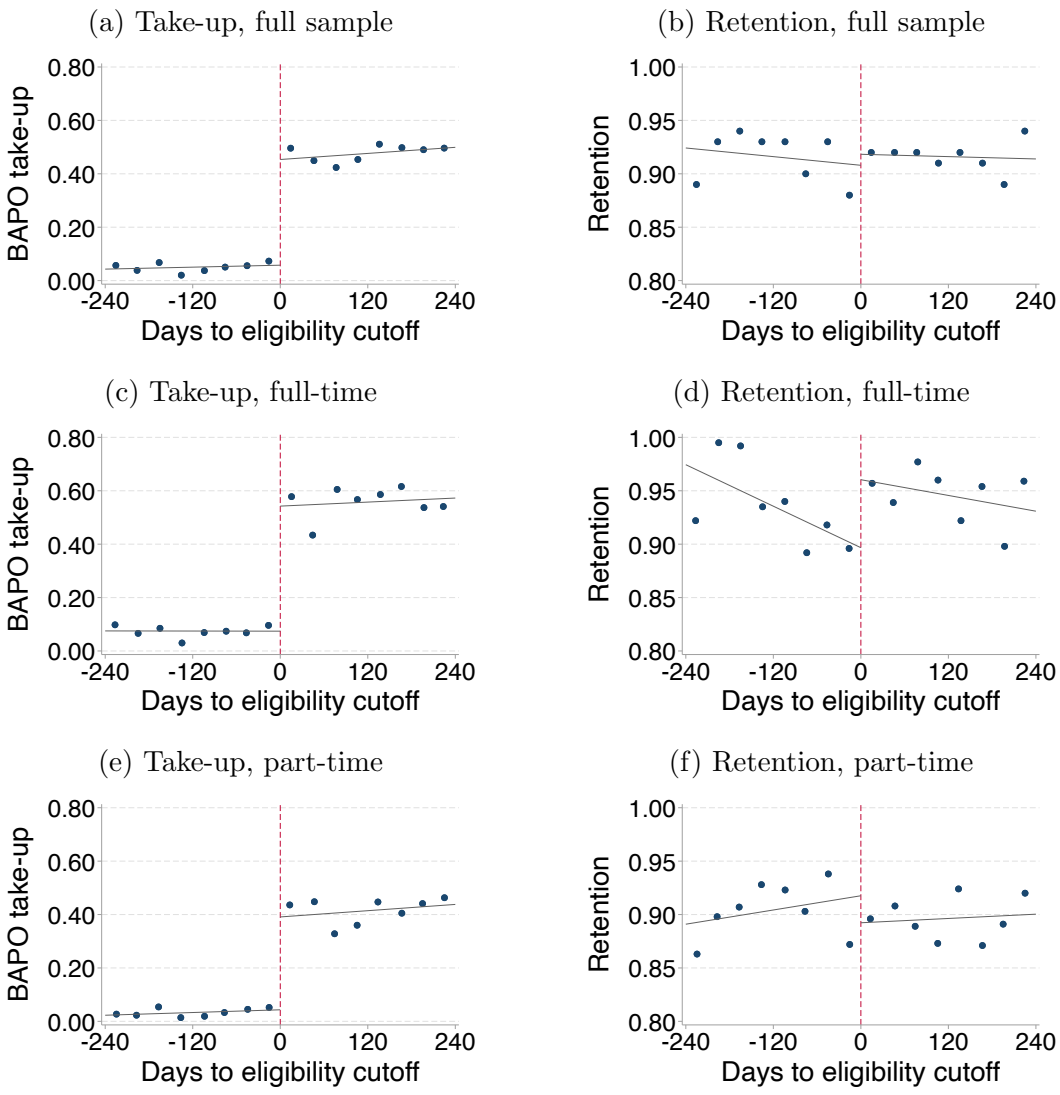
From 2016 onward, I observe the leave share, defined as subsidized leave FTE relative to contractual FTE. Eligibility increases this share by 0.039 overall, 0.047 among full-time, and 0.033 among part-time teachers. Scaling these estimates by the corresponding take-up effects yields complier leave shares of 0.10 for full-time and 0.096 for part-time teachers, close to the 0.10 entitlement cap. Participants in both groups therefore use nearly their full entitlement. Because leave share is observed only among teachers who remain in teaching, retention responses may affect the composition of the sample. Appendix D shows that this selection has only a modest effect on the estimated leave share.

5.2 Retention

Table 2 and Figure 2 report the effect of eligibility on retention in teaching. In the full sample, the estimate is 1.0 percentage point and not distinguishable from zero. This aggregate null masks differences by contract type. Among full-time teachers, eligibility increases retention by 6.1 percentage points, significant at the 5 percent level, relative to a control mean of 90 percent. Consistent with this estimate, the RD plot shows a jump at the cutoff against a declining gradient on either side. Among part-time teachers, the estimated effect is -2.7 percentage points and statistically insignificant. The difference between the two estimates, 8.8 percentage points, is itself significant at the 5 percent level. Taken together, these results suggest that the subsidy primarily retains teachers for whom it relaxes a constraint on hours.

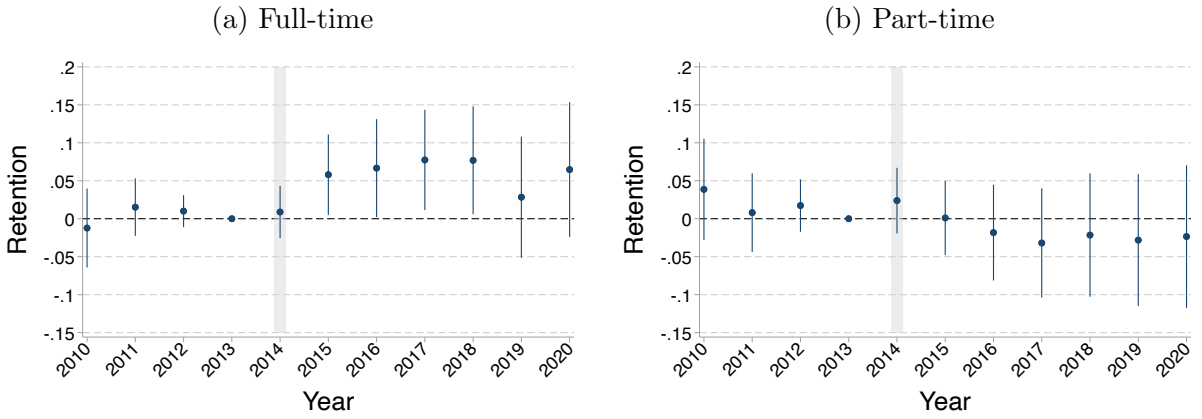
Year-by-year estimates show that the full-time effect appears in the first post-reform year and persists through 2018, whereas the part-time estimates are insignificant throughout (Figure 3). The gap does not close once the control group becomes eligible at 57: pooling 2015–2024, by which point the treatment cohort has reached 62, the full-time estimate is 6.3 percentage points, significant at the 10 percent level, and the part-time estimate is -3.7 points and insignificant. Teachers lost during the five-year deferral are therefore not recovered later. The full-time response reflects durable rather than tempo-

Figure 2: Effect of eligibility on BAPO take-up and retention in teaching



Notes: Panels (a), (c), and (e) plot BAPO take-up and panels (b), (d), and (f) plot retention in teaching against date of birth relative to the 1 August 1962 cutoff, for the full sample (top row) and separately for full-time (middle row) and part-time teachers (bottom row). Positive values denote teachers born before the cutoff, who are eligible from age 52. Both outcomes are pooled over 2015–2018, with evenly spaced 30-day bins and separate linear fits on each side; the bandwidth is 240 days with a uniform kernel. Take-up is an indicator for participating in any partially paid leave arrangement during the year; retention is an indicator for appearing in the teacher register as teaching staff in October. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

Figure 3: Year-by-year effect of eligibility on retention in teaching, 2010–2020



Notes: Each panel plots year-by-year RD estimates of equation (1) with 95 percent confidence intervals based on standard errors clustered at the teacher level. The shaded band marks 2014, excluded because the treatment group is only partially exposed in that year. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

rary exits: treating a teacher as having left only if she never reappears in the register yields an estimate of 0.060, equal to the main one (Table A.6), so the retention effect consists of teachers who would otherwise have left for good. For part-time teachers the permanent-exit estimate is -0.049, significant at the 10 percent level and somewhat more negative than the main -0.027, although the two are not statistically distinguishable.

Columns (2)–(4) of Table A.7 decompose the full-time retention effect by destination. Eligibility reduces the probability of holding a job outside education by 2.3 percentage points and raises the probability of being employed in any sector by 2.5 percentage points, both significant at the 10 percent level; the effect on moves into support-staff positions is small and insignificant. Taken together, about 40 percent of the retained teachers would otherwise have left employment, and most of the rest would have moved to jobs outside education.

Because take-up is incomplete, the estimates are intent-to-treat effects (ITT). Dividing the full-time estimate by the first stage of 0.47 gives an implied effect of about 13 percentage points among compliers. Reading this ratio as a local average treatment effect (LATE) requires that eligibility affect retention only through program use. That restriction may fail if eligibility is valuable in itself, for instance because the option to reduce hours later keeps a teacher in the job even if she never exercises it, in which case the

Table 2: Effect of eligibility on BAPO take-up, leave share, and retention

	BAPO take-up (1)	Leave share (2)	Retention (3)
Panel A: All			
RD estimate	0.395***	0.039***	0.010
	(0.031)	(0.003)	(0.020)
Control mean	0.059	0.000	0.912
Observations	8,747	5,921	8,747
Panel B: Full-time			
RD estimate	0.471***	0.047***	0.061**
	(0.049)	(0.005)	(0.028)
Control mean	0.078	0.001	0.900
Observations	3,740	2,595	3,740
Panel C: Part-time			
RD estimate	0.344***	0.033***	−0.027
	(0.038)	(0.004)	(0.028)
Control mean	0.043	0.001	0.921
Observations	5,007	3,326	5,007

Notes: The table reports RD estimates of equation (1). BAPO take-up is an indicator for participating in any partially paid leave arrangement at any point during the year. Retention is an indicator for appearing in the teacher register as teaching staff in October of the year. Both are pooled over 2015–2018. Leave share is the ratio of subsidized leave FTE to contractual FTE, pooled over 2016–2018, the period for which leave hours are recorded, and conditional on appearing in the teacher register. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

scaled estimate overstates the effect of participation.

Where the response occurs Because the Statistics Netherlands definition classifies teachers with contracts below 0.95 FTE as part-time, this group may combine teachers with substantially different levels of attachment to teaching. Table A.3 therefore examines heterogeneity by contract size, using mutually exclusive bins and progressively lower thresholds for full-time status. The estimates on retention differ across contract-size groups: 1.4 percentage points for contracts of 0.71–0.94 FTE, compared with −6.8 percentage points for contracts of 0.70 FTE or less. The negative estimate for teachers with the smallest contracts is consistent with the possibility that eligibility facilitates exits among teachers who are already weakly attached to teaching. For such teachers, the subsidy may relax the financial constraint associated with reducing their teaching hours or leaving teaching altogether, potentially facilitating transitions into self-employment, other jobs, or out of the labor force (Börsch-Supan et al., 2018). The estimate is too imprecise, however, to draw firm conclusions about this mechanism.

Differences in gender composition are unlikely to explain the pattern. Although women make up 73 percent of part-time teachers but only 31 percent of full-time teachers, the full-time effect is similar for men and women (Table A.8).

5.3 Total teaching supplied

Eligibility affects teaching supplied through two opposing margins. Teachers who use the leave teach fewer hours, while among full-time teachers, eligibility increases the probability of remaining in teaching.[24] Whether the program increases or reduces teaching supplied therefore depends on which response dominates. To assess this, I estimate the effect on total teaching supplied, coding teachers who are no longer in a teaching position as supplying zero hours. Because leave hours are recorded only from 2016 onward, the analysis covers 2016–2018.

In the full sample, the estimated effect is −0.032 FTE and statistically insignificant (Table 3; Figure A.6). The confidence interval rules out an increase larger than 0.015 FTE, equivalent to 2 percent of the control mean. Thus, there is no evidence that the program increases teaching time in the population as a whole.

The pooled estimate again masks two distinct responses. Among part-time teachers, eligibility reduces teaching supplied by 0.067 FTE, a statistically significant effect at the 5

[24] Appendix D discusses hours conditional on remaining in teaching. Because eligibility affects who remains, these conditional estimates are subject to selection and I treat them as descriptive.

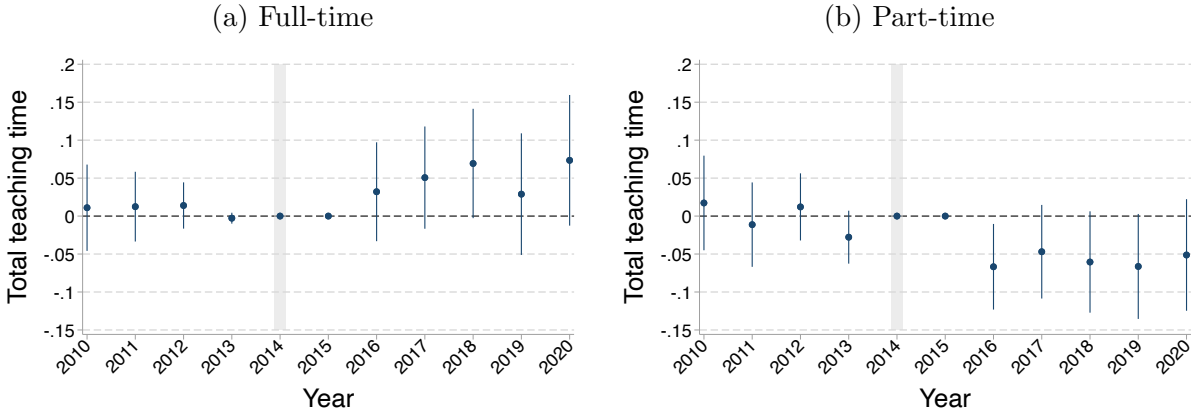
Table 3: Effect of eligibility on total teaching supplied, by contract type

	All	Full-time	Part-time
	(1)	(2)	(3)
RD estimate	−0.032	0.040	−0.067**
	(0.024)	(0.032)	(0.028)
Control mean	0.764	0.863	0.680
Observations	6,524	2,787	3,737

Notes: RD estimates of equation (1), pooled over 2016–2018, the period for which leave hours are recorded. The outcome is total teaching supplied in FTE: contractual FTE net of subsidized leave, with teachers no longer in the teacher register coded as zero. Full-time is ≥ 0.95 FTE in 2013. Bandwidth of 240 days, local linear specification, uniform kernel. Standard errors, clustered at the teacher level, in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

percent level. Because there is no corresponding retention response, the reduction in hours translates directly into lower teaching supplied. Among full-time teachers, by contrast, the estimated effect is 0.040 FTE and statistically insignificant; the confidence interval rules out a net reduction larger than 0.023 FTE, about 3 percent of the control mean. For full-time teachers, the additional hours supplied by retained teachers therefore largely offset the hours released through leave. Figure 4 reports the dynamic estimates separately for full-time and part-time teachers. The estimates are consistent with an increase in teaching supplied among full-time teachers, although the effects are not statistically significant, while teaching supplied declines among part-time teachers following access to the program. Both patterns persist after the control group becomes eligible: pooling 2016–2024, the full-time estimate is 0.056 FTE and insignificant, and the part-time estimate is −0.064 FTE, significant at the 5 percent level (Table A.9). The hours lost among part-time teachers are therefore not recovered once both groups have access to the program.

Figure 4: Year-by-year effect of eligibility on total teaching time, 2010–2020



Notes: Each panel plots year-by-year RD estimates of equation (1) with 95 percent confidence intervals based on standard errors clustered at the teacher level. Leave hours are recorded only from 2016; before 2016, total teaching time is measured as contractual FTE, with teachers no longer in the teacher register coded as zero. The shaded band marks 2014, excluded because the treatment group is only partially exposed in that year. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

5.4 Health Outcomes

Eligibility reduces hours in both groups but raises retention only among full-time teachers. Among part-time teachers, the health estimates therefore largely isolate the effect of a lighter workload. Among full-time teachers, they combine fewer hours with longer exposure to a demanding job. Because health is observed for all teachers in the 2013 sample regardless of subsequent employment, the estimates are not subject to selection from leaving teaching. Table 4 reports the estimates, and Figure A.7 shows the RD plots.

In the full sample, eligibility reduces pain-medication use by 5.7 percentage points, relative to a control mean of 21 percent, significant at the 1 percent level, but has no detectable effect on antidepressant use or high healthcare expenditure. These two estimates are imprecise, however, and cannot rule out sizable effects in either direction. The antidepressant estimate remains insignificant when mental health is measured more broadly as any use of generalist basic mental healthcare (Table A.11). There is also an increase of 5.2 percentage points in the probability of high general-practitioner expenditure, which is consistent with eligible teachers having more time to seek routine care rather than with worse health, given the decline in pain medication and the absence of any detectable change in total expenditure.

Among part-time teachers, eligibility reduces pain-medication use by 6.6 percentage points, significant at the 5 percent level, with no detectable effect on antidepressant use or

Table 4: Effect of eligibility on health outcomes, by contract type

	Antidepressants	Pain medication	Healthcare expenditure (>p50)
	(1)	(2)	(3)
Panel A: All			
RD estimate	0.007	−0.057***	0.003
	(0.019)	(0.022)	(0.031)
Control mean	0.064	0.212	0.560
Observations	8,747	8,747	8,747
Panel B: Full-time			
RD estimate	−0.003	−0.041	−0.001
	(0.025)	(0.035)	(0.047)
Control mean	0.053	0.240	0.545
Observations	3,740	3,740	3,740
Panel C: Part-time			
RD estimate	0.012	−0.066**	0.004
	(0.028)	(0.027)	(0.041)
Control mean	0.074	0.192	0.565
Observations	5,007	5,007	5,007

Notes: RD estimates of equation (1), pooled over 2015–2018. Antidepressants and Pain medication are indicators for any reimbursed prescription in the year (ATC N06A; N02A, N02B, or M01A). Healthcare expenditure (>p50) indicates total expenditure above the 2013 sample median. Outcomes are observed for all teachers in the 2013 sample, regardless of whether they remain in teaching. Full-time is ≥ 0.95 FTE in 2013. Bandwidth of 240 days, local linear specification, uniform kernel. Standard errors, clustered at the teacher level, in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

high healthcare expenditure. Because this group's retention does not show any significant change, the estimate more closely reflects the effect of a lighter workload and indicates improved physical health.

Among full-time teachers, none of the three outcomes changes significantly. The estimate for pain-medication use is −4.1 percentage points, similar in magnitude to the part-time estimate but imprecise. The confidence intervals rule out increases larger than 2.8 percentage points in pain-medication use, 4.6 in antidepressant use, and 9.1 in high healthcare expenditure. The estimates therefore exclude large increases in pain-medication use but not moderate deteriorations in mental health or healthcare use. The longer careers induced among full-time teachers are not accompanied by detectable harm on the out-

comes examined. Because eligibility both reduces their hours and extends their careers, however, the estimate reflects the net of the two, and the data cannot establish whether the additional years offset the health gain from fewer hours.

5.5 Robustness

Bandwidth sensitivity Figures A.8–A.10 report estimates of equation (1) at bandwidths from 90 to 330 days. Take-up is stable throughout. Among full-time teachers, retention loses significance only at the narrowest bandwidths, with an essentially unchanged point estimate, which reflects lower precision rather than instability. Among part-time teachers, the retention estimate moves toward zero and becomes more variable at narrow bandwidths. Teaching supplied declines somewhat at narrow bandwidths among full-time teachers but is insignificant throughout; among part-time teachers, it loses significance at narrow bandwidths but keeps its magnitude. Health estimates are generally stable, except that the estimates for pain-medication use and healthcare expenditure among part-time teachers shrink in magnitude at wider bandwidths.

Placebo cutoffs As a falsification test, I re-estimate equation (1) at 25 placebo cutoffs spaced one month apart, from May 1963 to May 1965, at which eligibility does not change (Cattaneo et al., 2020b).[25] The rank of the true estimate among all 26 estimates, counted in the direction of its sign, divided by 26 gives a one-sided randomization-inference p-value (Ganong and Jäger, 2018). Table A.12 reports the results. Take-up in all three samples, retention among full-time teachers, and teaching supplied among part-time teachers each rank first ($p = 1/26 \approx 0.04$, the smallest attainable value). The decline in pain-medication use ranks first in the full sample, third among part-time teachers, and fifth among full-time teachers, while antidepressant use and healthcare expenditure rank low in every sample.

Specification The main estimates are robust to how the sample is constructed, the set of controls, and the estimation method. Imputing the 193 person-years that appear in the employer–employee data under an education-sector employer but not in the teacher register as teaching employment, which keeps the panel balanced at 8,940 person-years, moves every coefficient by less than one standard error (Table A.2). Adding baseline covariates leaves take-up, full-time retention and the health estimates essentially unchanged, and part-time retention moves from -0.027 to -0.017, still insignificant (Table A.13).

[25] Day of birth is available only through 1965. Starting nine months after the true cutoff ensures that no placebo window contains it.

The main change is in part-time teaching supplied, which shrinks from −0.067 to −0.039 and loses significance, plausibly because the 2013 contract absorbs a small, insignificant difference in contract size at the cutoff within this group (Table A.4). Local randomization inference within ±30, ±60 and ±90 days of the cutoff (Cattaneo et al., 2016) gives the same picture: take-up jumps in every window, full-time retention is stable at 0.04–0.06 and significant at ±90 days, and part-time retention is close to zero throughout (Table A.14). With MSE-optimal bandwidths selected separately for each outcome and sample, a triangular kernel and robust bias-corrected inference (Calonico et al., 2014), the point estimates stay close to the main ones and their significance is unchanged, except that full-time teaching supplied becomes significant at the 10 percent level (Table A.15). A second-order polynomial on each side of the cutoff widens the confidence intervals but preserves the pattern: take-up remains large and precise, full-time retention is 0.083 (significant at the 10 percent level), and the pain-medication estimates keep their sign and magnitude (Table A.16). Dropping teachers born within one month of the cutoff leaves the full-time coefficients essentially unchanged, with retention at 0.058 against 0.061 in the main specification, although it loses significance (Table A.17). For part-time teachers, the donut keeps the magnitude of the pain-medication and teaching-supplied estimates but not their significance. This loss of precision is expected, because the donut excludes the teachers closest to the cutoff, who are the most informative. Across all specifications, part-time retention is never significantly positive, ranging from −0.059 in the donut (significant at the 10 percent level) to 0.030 with a quadratic.

No disappointment response in the control group Teachers born just after the cutoff learn that a few months of birth date cost them five years of eligibility. Such a narrow loss can itself trigger exit or worse mental health, which would confound the RD comparison (De Grip et al., 2012). Any such response should be strongest for the cohort that missed most narrowly and fade with distance from the cutoff, since later cohorts were excluded under the same rules but were still years from turning 52 when the reform was announced. I therefore compare the RD control cohort (born August 1962 to March 1963) with cohorts born one and two years later in a difference-in-differences design, using the most distant cohort as reference.[26] A disappointment effect would make the Cohort 1 interaction negative for retention, positive for medication use, and larger in absolute value than the Cohort 2 interaction. Because the cohorts differ in age, the interactions may also

[26]The sample consists of teachers employed in secondary education in 2013 in the three cohorts, observed over 2011–2018, with cohort and year fixed effects; Post equals one for 2015–2018.

absorb differences in the age profile of the outcomes, so the exercise is suggestive rather than a formal test.

Table A.18 shows no such response. Among full-time teachers, the Cohort 1 retention interaction is small and insignificant (−0.010). In the pooled sample it is −0.019 and significant at the 10 percent level, but this reflects part-time teachers. For them, the interactions (−0.026 and −0.016) decline roughly in proportion to the age gap, with no additional drop for the cohort closest to the cutoff, consistent with the age profile of exits rather than disappointment. A disappointment response among part-time teachers would in any case make their true effect more negative, reinforcing the conclusion that eligibility does not retain them. Antidepressant use shows no response in any group, and the pain-medication interactions are negative or close to zero, except for an insignificant positive Cohort 1 coefficient among full-time teachers.

6 Mechanisms

Eligibility may retain full-time teachers by relieving strain or by providing leave that teachers value as leisure. To distinguish between these channels, I compare full-time teachers in the lowest salary scale (LB) with full-time administrative and educational-support staff. Both groups are entitled to the leave, at a co-payment that is slightly lower for support staff (40 rather than 50 percent of forgone salary), but support staff report substantially less strain.[27]

Restricting teachers to scale LB makes the two groups more comparable on observables. In the linked survey sample (Table C.2), LB teachers earn about 20 percent more than support staff, and the groups have similar contract sizes, gender composition, and baseline health. Support staff are less likely to report high work pressure (37 versus 51 percent) and emotionally demanding work (10 versus 33 percent), and less likely to report that fewer hours would lead them to work longer (49 versus 67 percent). They are more likely to report physically demanding work (24 versus 14 percent).

Table 5 reports the estimates. Eligibility increases take-up by 39 percentage points among LB teachers and by 33 percentage points among support staff. It increases retention by 15.4 percentage points among LB teachers and by 0.7 percentage points among support staff. The LB estimate exceeds the 6.1 percentage points for all full-time teachers. This difference is consistent with the strain channel, since LB teachers are both closer to the

[27]The lower co-payment makes the leave worth slightly more to support staff. If the program operated through the value of the leave itself, they should therefore respond at least as strongly as teachers.

exit margin and more exposed to classroom strain. In the control group, 17 percent of LB teachers leave, against 10 percent of all full-time teachers. LB teachers are also concentrated in the pre-vocational track and rarely hold the coordinating duties that reduce classroom time (Bisschop and Heyma, 2016).

The comparison has one limitation. Only 3 percent of support staff in the control group leave, and the 95 percent confidence interval includes effects large enough to eliminate all of their exits. The comparison therefore does not establish that subsidized leave fails to retain low-strain workers with higher exit rates. It does show that take-up is similar across groups that differ in strain, whereas retention increases only in the group with both high strain and high exit risk. This pattern is consistent with the program operating through strain relief rather than through the value of leave as leisure.

7 Spillovers and Cost-Effectiveness

Whether subsidized workload reduction is worth its cost depends not only on how many teachers it retains, but also on how schools fill the hours that participants no longer teach. This section examines how schools cover the leave and then assesses the program's cost-effectiveness, accounting for both financial savings and gains in student achievement.

7.1 How Schools Cover the Leave

Schools can cover the lessons that teachers on leave no longer teach by expanding the contracts of incumbent teachers or by hiring. I exploit cross-school variation in exposure to the transitional arrangement to identify which margin they use.

Exposure A school's exposure is the 2013 FTE-weighted share of its teachers who kept early access under the transitional arrangement, that is, teachers born between August 1958 and July 1962:

$$\pi_s = \frac{\sum_{j \in s} \text{FTE}_{j,2013} \cdot \mathbf{1}[\text{born Aug 1958–Jul 1962}]}{\sum_{j \in s} \text{FTE}_{j,2013}}. \tag{2}$$

I use all four retained cohorts rather than births near the cutoff, which would generate little cross-school variation.[28] Because π_s is fixed by the 2013 workforce, it is predetermined with respect to any post-reform adjustment. Until August 2014, every teacher aged 52

[28] Teachers born before August 1958 were already 57 by 2015 and would have been eligible under either rule.

Table 5: Effect of eligibility on BAPO take-up and retention: lowest-scale teachers and support staff

	Teachers, scale LB		Support staff	
	BAPO take-up	Retention	BAPO take-up	Retention
	(1)	(2)	(3)	(4)
Panel A: All				
RD estimate	0.249***	0.004	0.272***	0.010
	(0.041)	(0.033)	(0.036)	(0.018)
Control mean	0.055	0.876	0.029	0.968
Observations	4,732	4,732	5,261	5,261
Panel B: Full-time				
RD estimate	0.393***	0.154***	0.331***	0.007
	(0.080)	(0.052)	(0.057)	(0.022)
Control mean	0.084	0.827	0.049	0.971
Observations	1,636	1,636	2,327	2,327
Panel C: Part-time				
RD estimate	0.204***	−0.064	0.231***	0.015
	(0.046)	(0.041)	(0.045)	(0.027)
Control mean	0.038	0.902	0.013	0.966
Observations	3,096	3,096	2,934	2,934

Notes: RD estimates of equation (1) for two samples: teachers in salary scale LB in 2013 (columns 1–2) and administrative and educational-support staff employed in secondary education in 2013 (columns 3–4). BAPO take-up is an indicator for participating in the program at any point during the year. Retention is an indicator for appearing in the personnel register in October in the same staff category: as a teacher in column 2 and as support staff in column 4. Both outcomes are pooled over 2015–2018. Full-time is ≥ 0.95 FTE in 2013. The bandwidth is 300 days, wider than the 240 days of the main tables to gain precision in these smaller samples; full-time estimates at 240 days are similar. Local linear specification, uniform kernel. Standard errors, clustered at the individual level, in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

or older could take the leave, so a school with many teachers born in 1958–62 was not distinctive. From August 2014 onward, these are precisely the teachers who keep early access while their slightly younger colleagues are deferred to age 57. I restrict the analysis to schools with at least 15 teachers in 2013.

Identification A school may have a high π_s simply because its workforce is older, and age structure shapes retirements and turnover independently of BAPO. Teachers born before August 1958, moreover, were eligible under both regimes. I therefore control for the 2013 FTE-weighted shares of the adjacent deferred cohort, π_s^{def} (born August 1962–July 1966), of always-eligible older teachers, π_s^{old} (born before August 1958), and of younger teachers, π_s^{young} (born 1967–75), each interacted with year indicators. The comparison is thus between schools with similar age structures that differ only in the share of teachers who retained early access.

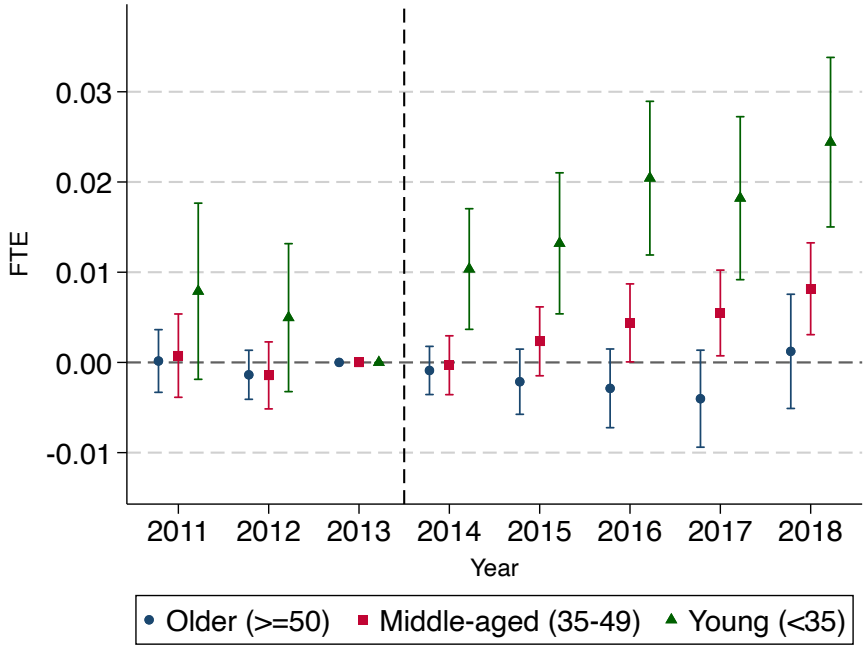
Coverage of withdrawn hours The outcome is the actual working FTE of incumbents who are themselves ineligible throughout the window, that is, born on or after August 1, 1962, so that any response reflects their colleagues' eligibility rather than their own. Each teacher is assigned to her 2013 school, and her FTE is set to zero if she leaves teaching. I estimate

$$y_{ist} = \alpha_i + \lambda_t + \sum_{t \neq 2013} \beta_t \left(\pi_s \times \mathbf{1}[\text{year} = t]\right) + \sum_{t \neq 2013} \left(\gamma_t \, \pi_s^{\text{def}} + \delta_t \, \pi_s^{\text{old}} + \eta_t \, \pi_s^{\text{young}}\right) \times \mathbf{1}[\text{year} = t] + \varepsilon_{ist}, \tag{3}$$

where α_i and λ_t are teacher and year fixed effects. Standard errors are clustered at the teacher level. Each β_t measures the differential change in incumbents' hours in more exposed schools relative to 2013, and the pre-2014 coefficients test for parallel trends. Figure 5 reports the estimates by age group in 2013, and table A.19 reports the pooled difference-in-differences estimates. The response is concentrated among young incumbents: a 10 percentage point increase in exposure raises the working FTE of teachers younger than 35 by 0.018 FTE, with the effect increasing gradually after the reform. The estimates for teachers aged 35–49 and 50 and older are 0.2 and 0.02 percentage points, respectively, and are not statistically significant. Young teachers also have the most room to expand, as their contracts are furthest below full time.

A second adjustment margin is hiring. I estimate the same event study at the school level, using the FTE of teachers newly employed at the school in year t, relative to the school's 2013 capacity. Figure A.11 shows no response to exposure after 2014. Hiring

Figure 5: Event-study estimates of school exposure on incumbent teachers' working hours, by age group



Notes: The figure plots event-study coefficients β_t from equation (3), separately for young (< 35), middle-aged (35–49), and older (≥ 50) incumbents, with age defined in 2013 and 2013 as the omitted reference year. Coefficients are scaled to a 10 percentage point increase in the exposure share π_s. The outcome is actual working FTE, set to zero for teachers no longer employed as teaching staff. The sample is restricted to schools with at least 15 teachers in 2013. Bars denote 95 percent confidence intervals based on standard errors clustered at the individual level.

therefore does not appear to be the margin through which schools adjust.

7.2 Cost-effectiveness

All figures are annual averages over the 2015–2018 eligibility window, in 2024 prices. Eligibility raises leave by 0.045 FTE per eligible full-time teacher and 0.021 FTE per eligible part-time teacher.[29] At a 65 percent subsidy and mean annual salaries of €51,300 and €34,200, the program costs about €1,500 per eligible full-time teacher and €660 per eligible part-time teacher per year.[30] Among full-time teachers, dividing by the 6.1 percentage-point retention effect gives about €25,000 per additional teacher-year retained,

[29]Leave share among stayers (0.047 and 0.033, Table 2) times the retention rate of eligible teachers (96 and 89 percent) times mean contract size (1.0 and 0.70). Leave hours are recorded only from 2016; I assume the 2015 leave share equals the 2016–2018 average.

[30]Salaries are October monthly salaries times twelve and exclude holiday allowance, year-end bonus, and employer contributions. All monetary figures scale proportionally with the salary base.

half a salary. Part-time teachers take up the leave without staying longer, so their subsidy buys fewer teaching hours rather than additional teacher-years.

The program also generates two savings for schools. First, the leave hours are taught mainly by younger incumbents, whom I assume earn 85 percent of senior pay. Second, each retained teacher-year avoids a vacancy, which costs about 5 percent of a salary more than keeping the senior, because part of the shortage is covered by agency staff at nearly twice the hourly cost (Appendix E.1). Per 100 eligible teachers, the subsidy costs about €103,000 per year and these savings recover about €31,000, leaving a net cost of €72,000: €43,000 for full-time and €29,000 for part-time teachers (Table E.2). In money terms, the program therefore does not pay for itself. Restricting eligibility to full-time teachers would cut the subsidy by 37 percent while preserving the entire retention effect, lowering the net cost per retained teacher-year from about €28,000 to about €17,000, a third of a salary.

For a government seeking to fill vacancies, the relevant benchmark is the cost of an additional teacher-year from other policies. The Dutch lateral-entry subsidy pays schools up to €25,000 per entrant. In England, training bursaries cost about £20,000 and early-career retention payments about £35,000 per additional teacher-year (McLean et al., 2023). At about €25,000, BAPO is thus comparable in cost, but it buys an experienced teacher-year in the school where the vacancy would arise, rather than a novice or early-career one.

The main return of retention is not financial but educational, because vacancies are mostly covered by teachers not qualified in the subject. Valuing student achievement with the earnings estimates of Chetty et al. (2014b), each avoided vacancy-year is worth about €840,000, while covering the leave hours with less experienced colleagues costs about €180,000 per FTE (Appendix E). Per 100 eligible teachers, the program then yields a net benefit of about €1.5 million per year: +€1.8 million from full-time and −€0.25 million from part-time teachers. The return remains positive as long as achievement is worth more than about 5 percent of the Chetty value. The program thus pays off through student achievement rather than money, and only through full-time teachers.

8 Conclusion

Workers in demanding occupations often leave well before retirement age, and in public-service professions facing persistent shortages, these exits are difficult to absorb. This paper asks whether reducing workload at later career stages helps retain experienced

workers. I exploit a 2014 reform that raised the eligibility age for subsidized leave for Dutch secondary-school teachers from 52 to 57 for those born on or after 1 August 1962, and compare teachers born on either side of this cutoff in a regression discontinuity design.

Eligibility raises take-up by 40 percentage points but has no detectable effect on retention. This aggregate null masks heterogeneity by contract size, a proxy for whether teachers had already reduced their hours at their own expense. Among full-time teachers, eligibility raises retention by 6.1 percentage points, an effect that persists through 2024, and the retention gains largely offset the hours released through leave. Part-time teachers take up the leave without remaining in teaching longer, so their teaching supplied falls by about 10 percent. Eligibility also reduces pain-medication use, an effect that is statistically significant only among part-time teachers, for whom the estimate largely isolates the effect of fewer hours. Full-time support staff, who are entitled to the same leave but report less strain, take it up at similar rates without staying longer, suggesting that the program retains teachers by relieving strain rather than by subsidizing leisure.

These findings have a direct implication for policy design. Subsidized workload reduction is most effective when it targets workers for whom it relaxes a binding constraint on hours, which in this setting are full-time teachers. Part-time teachers, by contrast, are more likely to have already adjusted their hours at their own expense; for them, the program lowers teaching supply without improving retention. Although the program does not pay for itself in fiscal terms, it generates a positive net return once its effects on student achievement are taken into account, and this return accrues entirely through full-time teachers. Restricting eligibility to full-time teachers would therefore substantially lower program costs while preserving its benefits. Under such targeting, the cost per retained teacher-year would be comparable to that of lateral-entry subsidies, training bursaries, and early-career retention payments. Unlike these alternatives, however, the program retains an experienced teacher in the school where the vacancy would otherwise arise.

References

Adriaens, H. and den Uijl, M. (2025). Arbeidsmarktramingen onderwijs 2025. Report, Centerdata, Tilburg. Commissioned by the Dutch Ministry of Education, Culture and Science.

Ahn, T. (2016). Reduction of working time: Does it lead to a healthy lifestyle? *Health Economics*, 25(8):969–983.

Albanese, A., Cockx, B., and Thuy, Y. (2020). Working time reductions at the end of the career: Do they prolong the time spent in employment? *Empirical Economics*, 59(1):99–141.

Algemene Rekenkamer (2014). Bekostiging voortgezet onderwijs. Technical report, Algemene Rekenkamer, Den Haag.

Algemene Rekenkamer (2024). Focus op inhuur van docenten. Focusonderzoek, Algemene Rekenkamer, The Hague.

Altonji, J. G. and Paxson, C. H. (1988). Labor supply preferences, hours constraints, and hours-wage trade-offs. *Journal of Labor Economics*, 6(2):254–276.

Altonji, J. G. and Paxson, C. H. (1992). Labor supply, hours constraints, and job mobility. *Journal of Human Resources*, 27(2):256–278.

Ameriks, J., Briggs, J., Caplin, A., Lee, M., Shapiro, M. D., and Tonetti, C. (2020). Older Americans would work longer if jobs were flexible. *American Economic Journal: Macroeconomics*, 12(1):174–209.

Bensnes, S., Hernæs, Ø. M., and King, M.-E. M. (2026). No payoff from time off? Mandated paid vacation and late-career employment. *Journal of Economic Behavior & Organization*, 245:107526.

Berniell, I. and Bietenbeck, J. (2020). The effect of working hours on health. *Economics & Human Biology*, 39:100901.

Biasi, B. (2021). The labor market for teachers under different pay schemes. *American Economic Journal: Economic Policy*, 13(3):1–45.

Bisschop, P. and Heyma, A. (2016). Op zoek naar LD-functies in het vmbo. Technical Report 2016-38, SEO Economisch Onderzoek, Amsterdam.

Blundell, R., Britton, J., Costa Dias, M., and French, E. (2023). The impact of health on labor supply near retirement. *Journal of Human Resources*, 58(1):282–334.

Börsch-Supan, A., Bucher-Koenen, T., Kutlu-Koc, V., and Goll, N. (2018). Dangerous flexibility – retirement reforms reconsidered. *Economic Policy*, 33(94):315–355.

Bratberg, E., Holmås, T. H., and Monstad, K. (2020). Health effects of reduced workload for older employees. *Health Economics*, 29(5):554–566.

Brown, K. (2013). The link between pensions and retirement timing: Lessons from California teachers. *Journal of Public Economics*, 98:1–14.

Buckles, K. S. and Hungerman, D. M. (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics*, 95(3):711–724.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, 14(4):909–946.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2020a). *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020b). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Cattaneo, M. D., Titiunik, R., and Vazquez-Bare, G. (2016). Inference in regression discontinuity designs under local randomization. *Stata Journal*, 16(2):331–367.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014a). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9):2593–2632.

Chetty, R., Friedman, J. N., and Rockoff, J. E. (2014b). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9):2633–2679.

Clotfelter, C. T., Glennie, E., Ladd, H. F., and Vigdor, J. L. (2008). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. *Journal of Public Economics*, 92(5–6):1352–1370.

Clotfelter, C. T., Ladd, H. F., and Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, 45(3):655–681.

Cygan-Rehm, K. and Wunder, C. (2018). Do working hours affect health? evidence from statutory workweek regulations in Germany. *Labour Economics*, 53:162–171.

De Grip, A., Lindeboom, M., and Montizaan, R. (2012). Shattered dreams: The effects of changing the pension system late in the game. *Economic Journal*, 122(559):1,–25.

Erick, P. N. and Smith, D. R. (2011). A systematic review of musculoskeletal disorders among school teachers. *BMC Musculoskeletal Disorders*, 12(260).

Euwals, R. (2001). Female labour supply, flexibility of working hours, and job mobility. *Economic Journal*, 111(471):120–134.

Falch, T. (2011). Teacher mobility responses to wage changes: Evidence from a quasi-natural experiment. *American Economic Review*, 101(3):460–465.

Filer, R. K. and Petri, P. A. (1988). A job-characteristics theory of retirement. *Review of Economics and Statistics*, 70(1):123–129.

Fouarge, D. and Baaijens, C. (2009). Job mobility and hours of work: The effect of Dutch legislation. ROA Research Memorandum 2009/4, Research Centre for Education and the Labour Market (ROA), Maastricht University.

Ganong, P. and Jäger, S. (2018). A permutation test for the regression kink design. *Journal of the American Statistical Association*, 113(522):494–504.

Gielen, A. C. (2009). Working hours flexibility and older workers' labor supply. *Oxford Economic Papers*, 61(2):240–274.

Graf, N., Hofer, H., and Winter-Ebmer, R. (2011). Labor supply effects of a subsidized old-age part-time scheme in Austria. *Journal for Labour Market Research (Zeitschrift für ArbeitsmarktForschung)*, 44(3):217–229.

Hendricks, M. (2014). Does it pay to pay teachers more? evidence from Texas. *Journal of Public Economics*, 109:45–56.

Herrmann, M. A. and Rockoff, J. E. (2012). Worker absence and productivity: Evidence from teaching. *Journal of Labor Economics*, 30(4):749–782.

Huber, C. A., Szucs, T. D., Rapold, R., and Reich, O. (2013). Identifying patients with chronic conditions using pharmacy data in Switzerland: an updated mapping approach to the classification of medications. *BMC Public Health*, 13:1030.

Lachowska, M., Mas, A., Saggio, R., and Woodbury, S. A. (2026). Work hours mismatch. *Econometrica*, 94(3):991–1025.

Leemans, A., van Bergen, K., Lubberman, J., Amsing, M., and Córvers, F. (2017). Personeel niet in loondienst (PNIL) in het po, vo en mbo. Report 16205, Regioplan Beleidsonderzoek, Amsterdam. Commissioned by the Dutch Ministry of Education, Culture and Science (OCW).

Leuven, E. (2006). Sick of work? the effect of work time reduction on the sickness absenteeism of teachers. Unpublished manuscript, University of Amsterdam (UvA).

McLean, D., Tang, S., and Worth, J. (2023). The impact of training bursaries on teacher recruitment and retention: An evaluation of impact and value for money. Technical report, National Foundation for Educational Research (NFER), Slough.

Ministerie van Onderwijs, Cultuur en Wetenschap (2024). Personeelstekorten voortgezet onderwijs, peildatum 1 oktober 2024. Technical report, Ministerie van OCW, Den Haag. Bijlage bij de Trendrapportage Arbeidsmarkt Leraren po, vo en mbo 2024.

OECD (2020). *TALIS 2018 Results (Volume II): Teachers and School Leaders as Valued Professionals*. TALIS. OECD Publishing, Paris.

OECD (2023). *Education at a Glance 2023: OECD Indicators*. Education at a Glance. OECD.

OECD (2025). The demands of teaching. In *Results from TALIS 2024: The State of Teaching*. OECD Publishing, Paris.

Ravesteijn, B., van Kippersluis, H., and van Doorslaer, E. (2018). The wear and tear on health: What is the role of occupation? *Health Economics*, 27(2):e69–e86.

Rivkin, S. G., Hanushek, E. A., and Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, 73(2):417–458.

Ruffini, K. (2024). Worker earnings, service quality, and firm profitability: Evidence from nursing homes and minimum wage reforms. *The Review of Economics and Statistics*, 106(6):1477–1494.

Rutten, A., Knoef, M., and van Vuuren, D. (2025). Employment effects of incentivized gradual retirement plans. *Work, Aging and Retirement*, 11(1):47–63.

Sánchez, R. (2017). Does a mandatory reduction of standard working hours improve employees' health status? *Industrial Relations: A Journal of Economy and Society*, 56(1):3–39.

Sims, S. and Benhenda, A. (2022). The effect of financial incentives on the retention of shortage-subject teachers: Evidence from England. CEPEO Working Paper 22-04, Centre for Education Policy and Equalising Opportunities (CEPEO), UCL.

TNO (2024). Factsheet fysieke arbeidsbelasting 2024. Factsheet, TNO, Leiden.

Tweede Kamer (2011). Motie van het lid Elias over het afbouwen van de BAPO-regeling. Kamerstukken II 2010/11, 27 923. Ingediend 26 april 2011, aangenomen 28 april 2011.

Van Casteren, W., Lodewick, J., Lommertzen, J., Luyten, E., and van Mensvoort, C. (2023). Vertrekredenen leraren en docenten in het po, vo en mbo. Technical report, ResearchNed, Nijmegen. Commissioned by the Dutch Ministry of Education, Culture and Science.

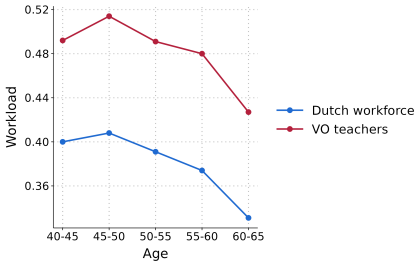
van den Heuvel, S. and de Vroome, E. (2023). Werkdruk in het onderwijs. Report TNO 2023 R12511, TNO, Leiden.

VO-raad (2014). CAO VO 2014–2015: Collectieve arbeidsovereenkomst voor het voortgezet onderwijs. Utrecht: VO-raad. Hoofdstuk 7: Levensfasebewust personeelsbeleid; art. 7.4: Overgangsregeling.

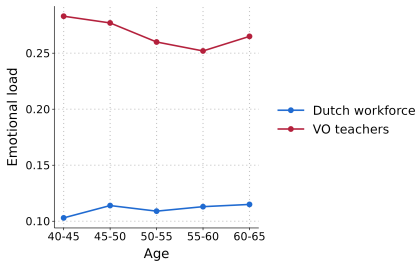
Voion (2026). Levensfasebewust personeelsbeleid in de praktijk: Onderzoek naar behoeften, knelpunten en ontwikkelkansen in het voortgezet onderwijs. Technical report, Voion, Arbeidsmarkt- en opleidingsfonds voortgezet onderwijs.

A Additional Tables and Figures

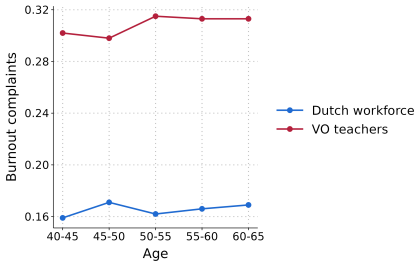
Figure A.1: Work-related strain by age group, secondary-school teachers versus the Dutch workforce



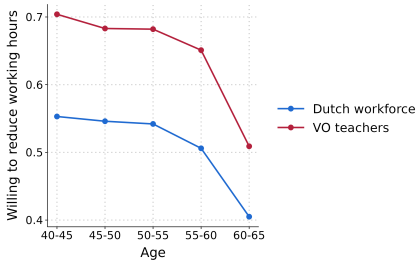
(a) High workload



(b) High emotional demands



(c) Burnout complaints



(d) Willing to reduce working hours

Notes: Each panel shows the share of employees reporting the outcome, by age group, for secondary-school (VO) teachers and for the Dutch workforce as a whole. Variable definitions are given in Appendix C. Source: Netherlands Working Conditions Survey (NEA), 2014–2024.

Table A.1: Effect of eligibility on attrition from the population register

	All	Full-time	Part-time
	(1)	(2)	(3)
RD estimate	0.010	0.009	0.011
	(0.010)	(0.012)	(0.014)
Control mean	0.009	0.007	0.011
Observations	9,076	3,852	5,224

Notes: The table reports RD estimates of equation (1), pooled over 2015–2018. The outcome is an indicator equal to one if the teacher ceases to be registered as a Dutch resident in the year. The sample consists of the 2,278 teachers of the main sample before the residency restriction described in Appendix B; teachers are not observed after they cease to be resident. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.2: Robustness: balanced panel

	RD estimate		
	All	Full-time	Part-time
	(1)	(2)	(3)
Panel A: Extensive margin			
BAPO take-up	0.394***	0.471***	0.342***
	(0.030)	(0.048)	(0.038)
Retention	0.004	0.060**	−0.036
	(0.020)	(0.028)	(0.028)
Observations	8,940	3,808	5,132
Panel B: Intensive margin			
Teaching supplied (FTE)	−0.039	0.039	−0.077***
	(0.024)	(0.032)	(0.028)
Observations	6,687	2,846	3,841
Panel C: Health			
Antidepressants	0.008	−0.005	0.016
	(0.019)	(0.025)	(0.027)
Pain medication	−0.054**	−0.041	−0.060**
	(0.021)	(0.035)	(0.027)
Healthcare expenditure (>p50)	0.000	−0.010	0.004
	(0.031)	(0.047)	(0.040)
Observations	8,940	3,808	5,132

Notes: The table reports RD estimates of equation (1) on the balanced panel, in which the 193 person-years observed in the employer–employee register under an education-sector employer but absent from the teacher register are imputed as teaching employment (Appendix B). Outcomes are defined as in Tables 2–4 and pooled over 2015–2018 (2016–2018 for Teaching supplied (FTE) in panel B). Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.3: Effect of eligibility on retention, by 2013 contract size

Panel A: Contract-size bins			
	≤ 0.70 FTE	0.71–0.94 FTE	≥ 0.95 FTE
	(1)	(2)	(3)
RD estimate	−0.068	0.014	0.061**
	(0.050)	(0.031)	(0.028)
Control mean	0.918	0.916	0.900
Observations	2,210	2,797	3,740
Panel B: Cumulative full-time thresholds			
	≥ 0.70 FTE	≥ 0.85 FTE	≥ 0.95 FTE
	(1)	(2)	(3)
RD estimate	0.045**	0.060**	0.061**
	(0.021)	(0.025)	(0.028)
Control mean	0.907	0.906	0.900
Observations	6,731	4,392	3,740

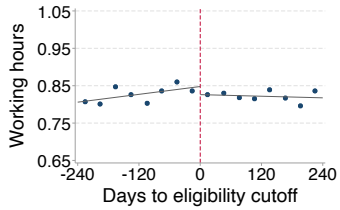
Notes: Each cell reports an RD estimate of equation (1) for retention in teaching, pooled over 2015–2018, on the subsample defined by 2013 contract size in the column heading. Panel A uses mutually exclusive bins; panel B uses cumulative thresholds, so column 3 of each panel reproduces the full-time estimate of Table 2. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.4: Covariate balance at the cutoff: predetermined characteristics and pre-reform outcomes, by contract type

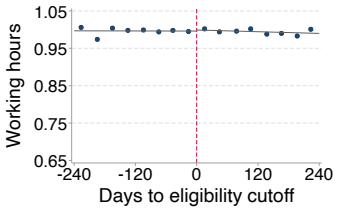
	All		Full-time		Part-time	
	RD est.	Std. err.	RD est.	Std. err.	RD est.	Std. err.
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Demographics and labor market						
Female	0.062	(0.042)	0.032	(0.058)	0.046	(0.051)
Immigrant background	−0.013	(0.032)	−0.013	(0.046)	−0.015	(0.044)
Contractual working hours	−0.021	(0.016)	0.000	(0.006)	−0.012	(0.018)
Monthly salary (euros)	−82.6	(109.1)	42.9	(92.7)	−16.1	(116.4)
Panel B: Health						
Antidepressants	0.012	(0.020)	−0.013	(0.023)	0.029	(0.030)
Pain medication	−0.015	(0.026)	−0.011	(0.040)	−0.019	(0.034)
Healthcare expenditure (>p50)	−0.024	(0.036)	−0.002	(0.053)	−0.049	(0.048)
Observations	4,389		1,877		2,512	

Notes: The table reports RD estimates of equation (1) at the August 1, 1962, cutoff using pre-reform data, pooled over 2012–2013. Female and Immigrant background are time invariant and are estimated on the same person-years as the other rows for comparability. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

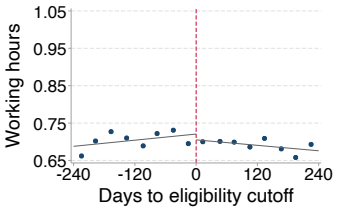
Figure A.2: Covariate balance: pre-reform employment and health outcomes



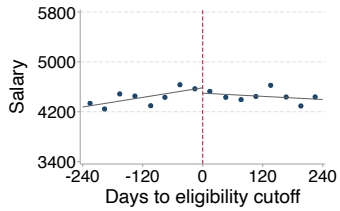
(a) Contractual working hours, full sample



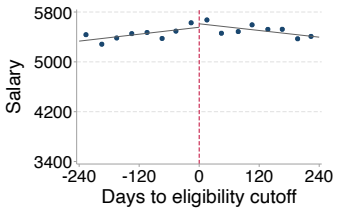
(b) Contractual working hours, full-time



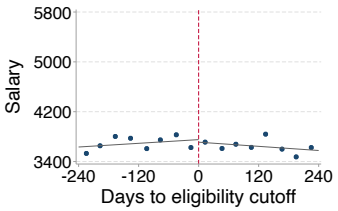
(c) Contractual working hours, part-time



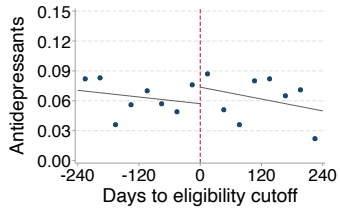
(d) Monthly salary, full sample



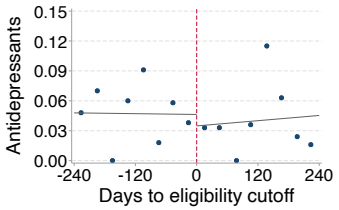
(e) Monthly salary, full-time



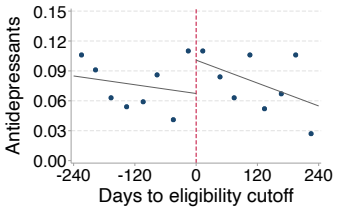
(f) Monthly salary, part-time



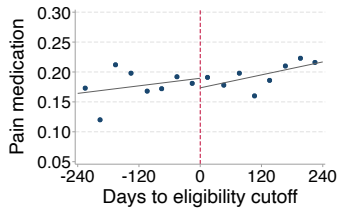
(g) Antidepressants, full sample



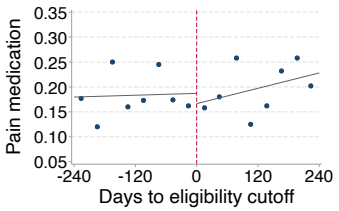
(h) Antidepressants, full-time



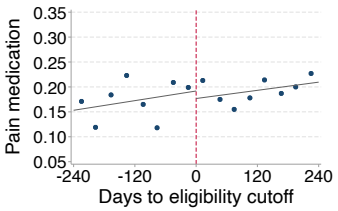
(i) Antidepressants, part-time



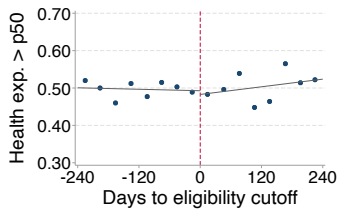
(j) Pain medication, full sample



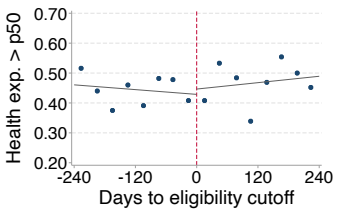
(k) Pain medication, full-time



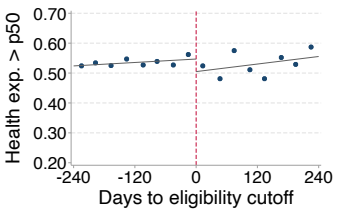
(l) Pain medication, part-time



(m) Healthcare expenditure (>p50), full sample



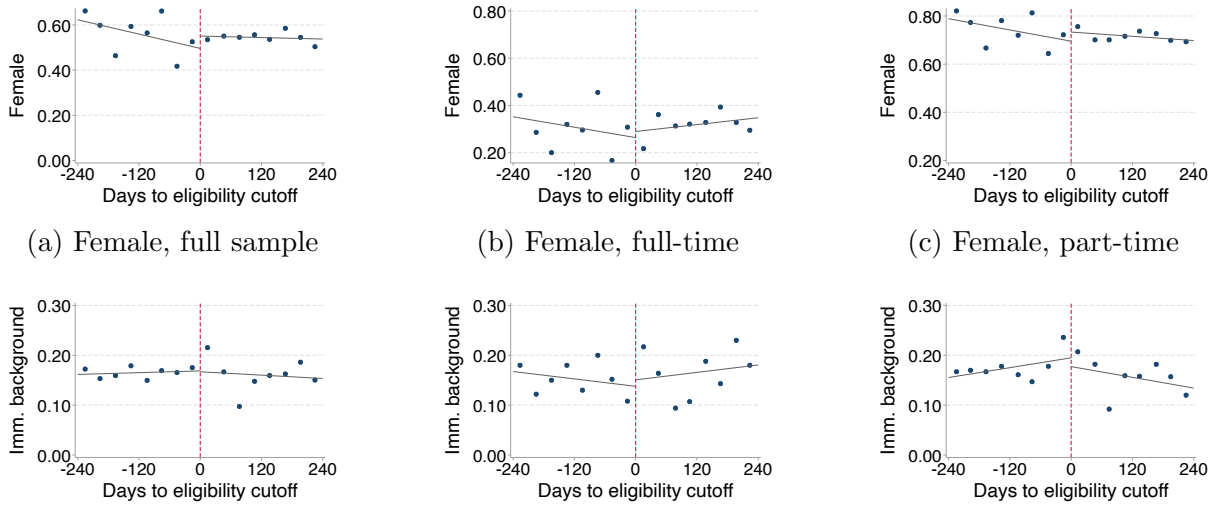
(n) Healthcare expenditure (>p50), full-time



(o) Healthcare expenditure (>p50), part-time

Notes: Each panel plots the pre-reform value of the variable indicated in the panel title against date of birth relative to the 1 August 1962 cutoff, for the full sample and separately for full- and part-time teachers, with evenly spaced 30-day bins and separate linear fits on each side; the bandwidth is 240 days with a uniform kernel. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

Figure A.3: Covariate balance: demographic characteristics



(a) Female, full sample

(b) Female, full-time

(c) Female, part-time

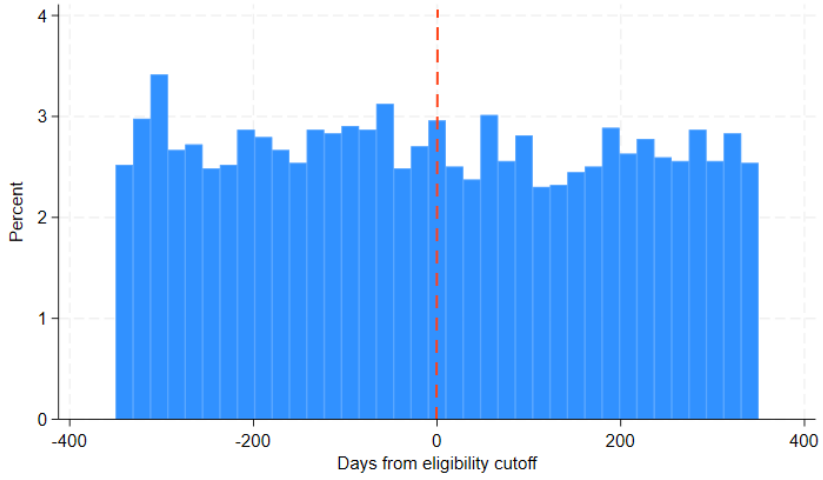
(d) Immigrant background, full sample

(e) Immigrant background, full-time

(f) Immigrant background, part-time

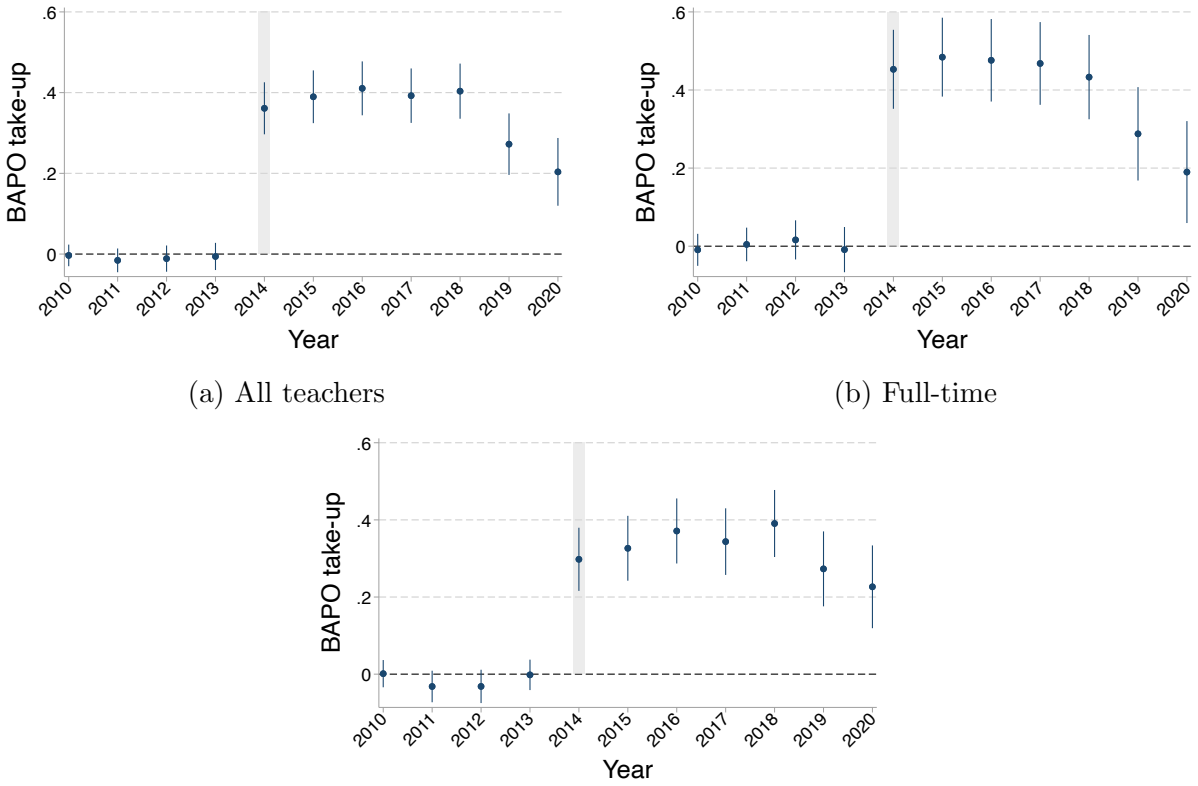
Notes: Each panel plots the predetermined covariate indicated in the panel title against date of birth relative to the 1 August 1962 cutoff, for the full sample and separately for full- and part-time teachers, with evenly spaced 30-day bins and separate linear fits on each side; the bandwidth is 240 days with a uniform kernel. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

Figure A.4: Manipulation test of the running variable



Notes: The figure plots the density of the running variable (days from the 1 August 1962 cutoff) for the analysis sample. The vertical dashed line marks the cutoff.

Figure A.5: Year-by-year effect of eligibility on BAPO take-up, 2010–2020



(a) All teachers

(b) Full-time

(c) Part-time

Notes: Each panel plots year-by-year RD estimates of equation (1) with 95 percent confidence intervals based on standard errors clustered at the teacher level. The shaded band marks 2014, excluded because the treatment group is only partially exposed in that year. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

Table A.5: Long-run effect of eligibility on retention and destination of exits, by contract type, 2015–2024

	Retention	Non-education	Support staff	Total employment
	(1)	(2)	(3)	(4)
Panel A: All				
RD estimate	0.005	0.004	0.008	0.012
	(0.024)	(0.010)	(0.012)	(0.017)
Control mean	0.830	0.030	0.024	0.890
Observations	21,689	21,689	21,689	21,689
Panel B: Full-time				
RD estimate	0.063*	−0.007	−0.016	0.036*
	(0.033)	(0.013)	(0.021)	(0.020)
Control mean	0.834	0.028	0.040	0.910
Observations	9,262	9,262	9,262	9,262
Panel C: Part-time				
RD estimate	−0.037	0.011	0.028**	−0.002
	(0.034)	(0.014)	(0.014)	(0.026)
Control mean	0.826	0.031	0.011	0.871
Observations	12,427	12,427	12,427	12,427
Panel D: Difference (full-time − part-time)				
RD estimate	0.100**	−0.017	−0.044*	0.038
	(0.047)	(0.019)	(0.025)	(0.033)
Observations	21,689	21,689	21,689	21,689

Notes: The table reports RD estimates of equation (1), pooled over 2015–2024, so that the window includes the years from 2019 onward in which the control cohort is also eligible. Outcomes are defined as in Table A.7. Panel D reports the difference between the full-time and part-time estimates. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.6: Effect of eligibility on permanent retention in teaching

	All	Full-time	Part-time
	(1)	(2)	(3)
RD estimate	−0.002	0.060**	−0.049*
	(0.018)	(0.024)	(0.025)
Control mean	0.960	0.936	0.978
Observations	8,747	3,740	5,007

Notes: The table reports RD estimates of equation (1), pooled over 2015–2018. The outcome is an absorbing measure of retention that treats a teacher as having exited in year t only if she never reappears in the teacher register thereafter. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.7: Effect of eligibility on retention and destination of exits, by contract type

	Retention	Non-education	Support staff	Total employment
	(1)	(2)	(3)	(4)
Panel A: All				
RD estimate	0.010	−0.015*	0.005	−0.004
	(0.020)	(0.009)	(0.011)	(0.013)
Control mean	0.912	0.026	0.017	0.960
Observations	8,747	8,747	8,747	8,747
Panel B: Full-time				
RD estimate	0.061**	−0.023*	−0.007	0.025*
	(0.028)	(0.012)	(0.020)	(0.013)
Control mean	0.900	0.028	0.019	0.963
Observations	3,740	3,740	3,740	3,740
Panel C: Part-time				
RD estimate	−0.027	−0.009	0.015	−0.024
	(0.028)	(0.012)	(0.012)	(0.021)
Control mean	0.921	0.024	0.007	0.960
Observations	5,007	5,007	5,007	5,007
Panel D: Difference (full-time − part-time)				
RD estimate	0.088**	−0.014	−0.022	0.048*
	(0.040)	(0.017)	(0.023)	(0.025)
Observations	8,747	8,747	8,747	8,747

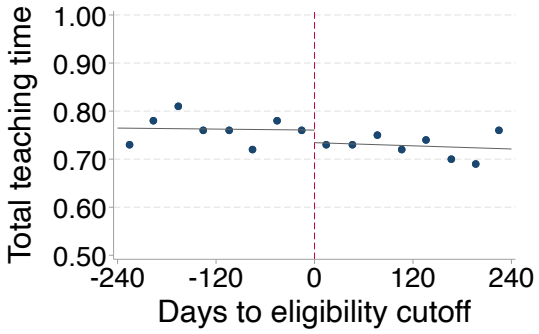
Notes: The table reports RD estimates of equation (1), pooled over 2015–2018. Retention is an indicator for appearing in the teacher register as teaching staff in October of the year. Columns 2–4 decompose destinations: Non-education is an indicator for holding a job outside the education sector, Support staff for remaining in education in an administrative or educational-support role, and Total employment for being employed in any sector; all are constructed from the teacher and employer–employee registers and measured in October. Panel D reports the difference between the full-time and part-time estimates. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.8: Effect of eligibility, by contract type and gender

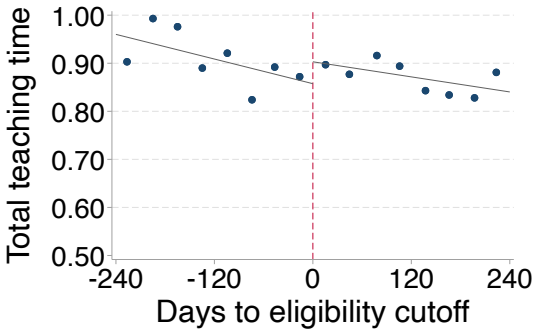
	Women			Men		
	All	Full-time	Part-time	All	Full-time	Part-time
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Extensive margin						
BAPO take-up	0.386***	0.458***	0.367***	0.408***	0.476***	0.282***
	(0.035)	(0.074)	(0.039)	(0.042)	(0.052)	(0.069)
Retention	−0.012	0.052	−0.031	0.026	0.053*	−0.021
	(0.025)	(0.043)	(0.030)	(0.025)	(0.031)	(0.043)
Observations	5,978	1,492	4,486	5,036	3,294	1,742
Panel B: Working hours						
Teaching supplied (FTE)	−0.045	0.058	−0.074**	−0.016	0.013	−0.027
	(0.028)	(0.050)	(0.029)	(0.030)	(0.035)	(0.049)
Observations	4,463	1,114	3,349	3,755	2,456	1,299
Panel C: Health						
Antidepressants	0.005	0.012	0.003	0.006	0.003	0.012
	(0.026)	(0.046)	(0.031)	(0.021)	(0.025)	(0.039)
Pain medication	−0.045*	−0.122*	−0.018	−0.036	−0.015	−0.069
	(0.027)	(0.064)	(0.029)	(0.027)	(0.035)	(0.044)
Healthcare expenditure (>p50)	−0.004	−0.130*	0.034	0.000	0.033	−0.058
	(0.037)	(0.073)	(0.042)	(0.042)	(0.051)	(0.076)
Observations	5,978	1,492	4,486	5,036	3,294	1,742

Notes: The table reports RD estimates of equation (1), separately by contract type and gender, pooled over 2015–2018 (2016–2018 for Teaching supplied (FTE) in panel B). Each cell corresponds to a separate regression. Outcomes are defined as in Tables 2–4; Teaching supplied (FTE) in panel B is unconditional on remaining in teaching. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is widened to 300 days around the August 1, 1962, cutoff to improve precision in the gender-specific subsamples, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

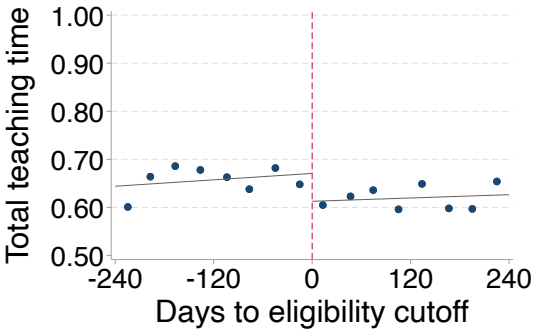
Figure A.6: Effect of eligibility on total teaching time



(a) All teachers



(b) Full-time



(c) Part-time

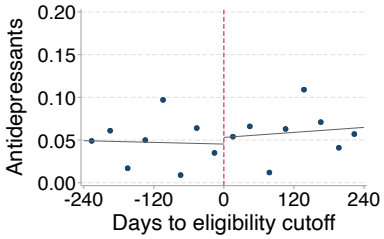
Notes: Each panel plots total teaching time against date of birth relative to the 1 August 1962 cutoff, pooled over 2016–2018, with evenly spaced 30-day bins and separate linear fits on each side; the bandwidth is 240 days with a uniform kernel. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

Table A.9: Long-run effect of eligibility on total teaching supplied, by contract type, 2016–2024

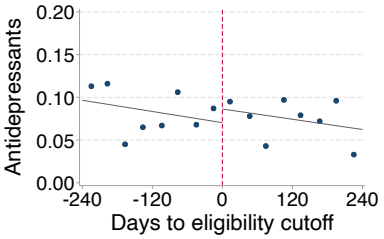
	All	Full-time	Part-time
	(1)	(2)	(3)
RD estimate	−0.021	0.056	−0.064**
	(0.024)	(0.034)	(0.029)
Control mean	0.670	0.758	0.590
Observations	19,446	8,295	11,151

Notes: RD estimates of equation (1), pooled over 2016–2024, so that the window includes the years from 2019 onward in which the control cohort is also eligible. The outcome is total teaching supplied, defined as in Table 3. Full-time is ≥ 0.95 FTE in 2013. Bandwidth of 240 days, local linear specification, uniform kernel. Standard errors, clustered at the teacher level, in parentheses. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

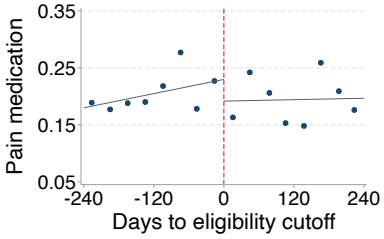
Figure A.7: RD plots for health outcomes, by contract type



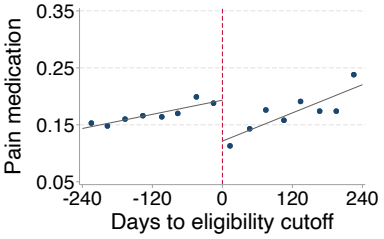
(a) Full-time: antidepressants



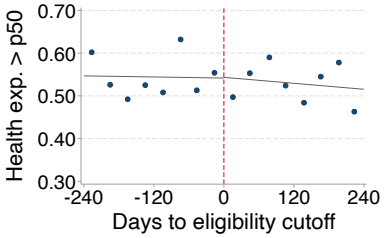
(b) Part-time: antidepressants



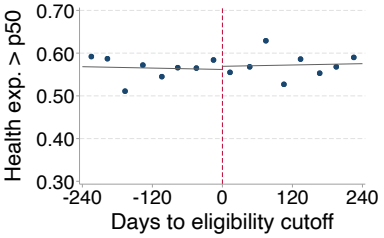
(c) Full-time: pain medication



(d) Part-time: pain medication



(e) Full-time: healthcare expenditure (>p50)



(f) Part-time: healthcare expenditure (>p50)

Notes: Each panel plots the health outcome indicated in the panel title against date of birth relative to the 1 August 1962 cutoff, pooled over 2015–2018, with evenly spaced 30-day bins and separate linear fits on each side; the bandwidth is 240 days with a uniform kernel. Outcomes are observed for all teachers in the 2013 sample and are therefore not conditional on remaining in teaching. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

Table A.10: Long-run effect of eligibility on health outcomes, by contract type, 2015–2024

	Antidepressants	Pain medication	Healthcare expenditure (>p50)
	(1)	(2)	(3)
Panel A: All			
RD estimate	−0.001	−0.034*	0.023
	(0.019)	(0.017)	(0.028)
Control mean	0.073	0.184	0.581
Observations	21,689	21,689	21,689
Panel B: Full-time			
RD estimate	−0.020	−0.027	0.003
	(0.025)	(0.027)	(0.042)
Control mean	0.070	0.197	0.576
Observations	9,262	9,262	9,262
Panel C: Part-time			
RD estimate	0.012	−0.037*	0.038
	(0.027)	(0.022)	(0.037)
Control mean	0.075	0.172	0.580
Observations	12,427	12,427	12,427

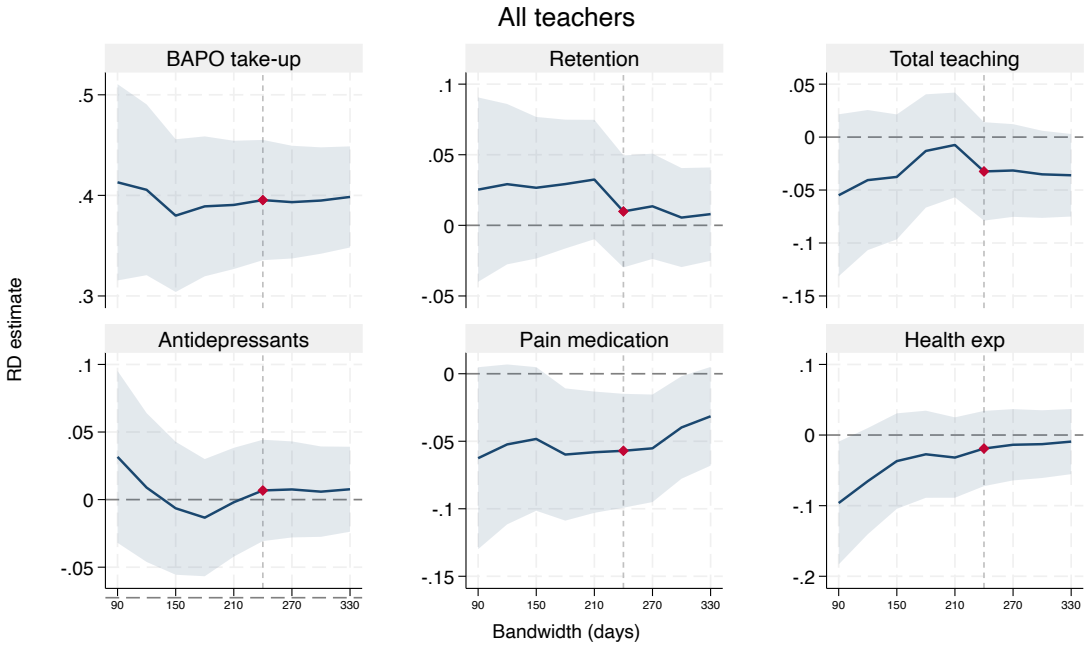
Notes: The table reports RD estimates of equation (1), pooled over 2015–2024, so that the window includes the years from 2019 onward in which the control cohort is also eligible. Outcomes are defined as in Table 4 and are not conditional on remaining in teaching. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.11: Effect of eligibility on additional health outcomes, by contract type

	Mental healthcare (broad) (1)	Mental healthcare expenditure (2)	GP expenditure (>p75) (3)
Panel A: All			
RD estimate	0.013	0.003	0.052**
	(0.020)	(0.007)	(0.026)
Control mean	0.083	0.025	0.276
Observations	8,747	8,747	8,747
Panel B: Full-time			
RD estimate	−0.005	−0.005	0.029
	(0.027)	(0.009)	(0.038)
Control mean	0.063	0.019	0.241
Observations	3,740	3,740	3,740
Panel C: Part-time			
RD estimate	0.024	0.008	0.065*
	(0.029)	(0.011)	(0.036)
Control mean	0.100	0.031	0.305
Observations	5,007	5,007	5,007

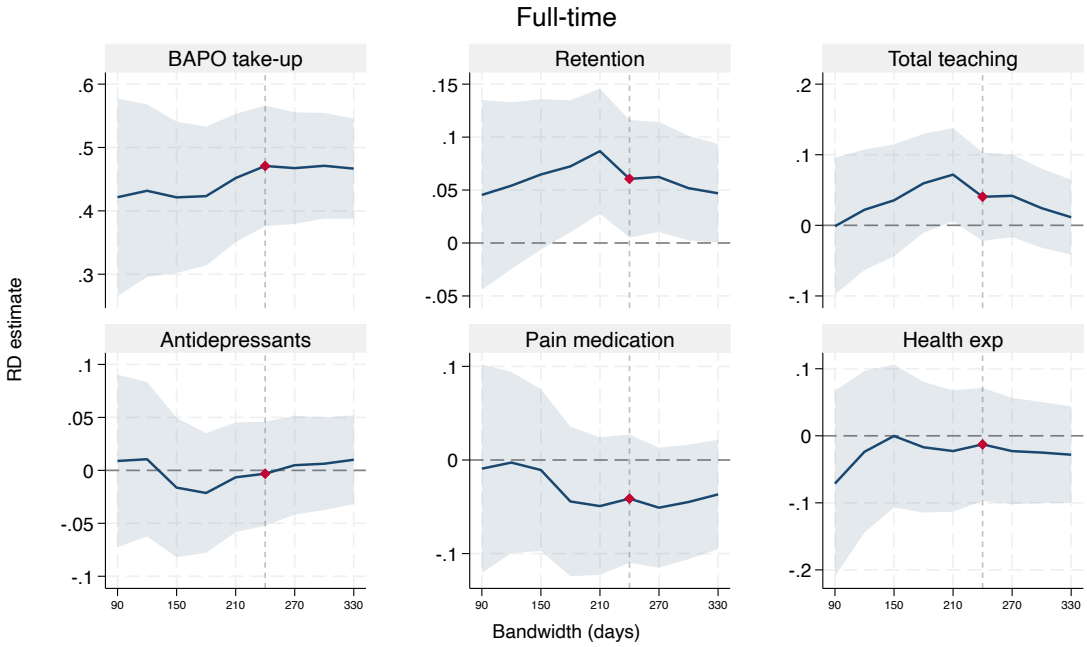
Notes: The table reports RD estimates of equation (1), pooled over 2015–2018. Mental healthcare (broad) is an indicator equal to one if the teacher fills an antidepressant prescription or has any reimbursed expenditure on generalist basic mental healthcare in the year. Mental healthcare expenditure is an indicator for any reimbursed generalist basic mental healthcare expenditure. GP expenditure (>p75) is an indicator for general-practitioner expenditure above the 75th percentile within the analysis sample. Outcomes are not conditional on remaining in teaching. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Figure A.8: Bandwidth sensitivity, all teachers



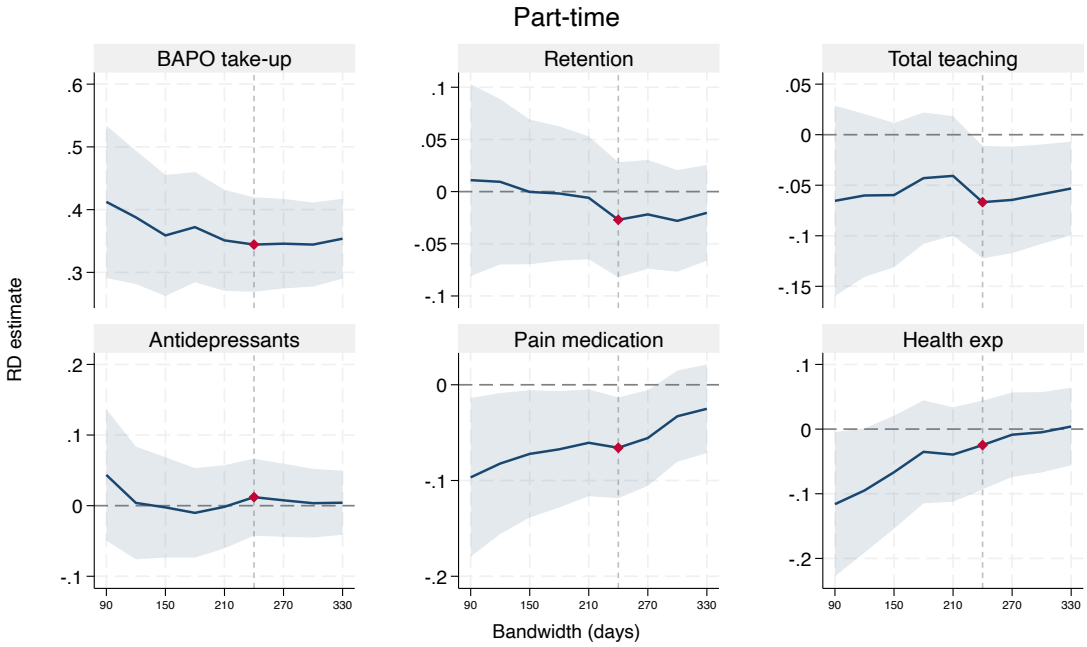
Notes: The figure plots RD estimates and 95 percent confidence intervals for the main outcomes across alternative bandwidths ranging from 90 to 330 days. Estimates correspond to equation (1) pooled over the post-reform period 2015–2018 (2016–2018 for the intensive-margin outcomes).

Figure A.9: Bandwidth sensitivity, full-time teachers



Notes: The figure plots RD estimates and 95 percent confidence intervals for the main outcomes across alternative bandwidths ranging from 90 to 330 days, for full-time teachers ($\geq$ 0.95 FTE in 2013). Estimates correspond to equation (1) pooled over the post-reform period 2015–2018 (2016–2018 for the intensive-margin outcomes).

Figure A.10: Bandwidth sensitivity, part-time teachers



Notes: The figure plots RD estimates and 95 percent confidence intervals for the main outcomes across alternative bandwidths ranging from 90 to 330 days, for part-time teachers (below 0.95 FTE in 2013). Estimates correspond to equation (1) pooled over the post-reform period 2015–2018 (2016–2018 for the intensive-margin outcomes).

Table A.12: Placebo-cutoff estimates and randomization-inference p-values

Outcome	True estimate	Placebo mean	Rank	p
Panel A: All teachers				
BAPO take-up	0.395	0.003	1	0.038
Retention	0.010	0.001	9	0.346
Total teaching supplied	−0.032	0.003	2	0.077
Antidepressants	0.007	0.000	9	0.346
Pain medication	−0.057	0.003	1	0.038
Healthcare expenditure (>p50)	0.003	0.006	13	0.500
Panel B: Full-time teachers				
BAPO take-up	0.471	0.004	1	0.038
Retention	0.061	−0.002	1	0.038
Total teaching supplied	0.040	0.002	6	0.231
Antidepressants	−0.003	0.000	13	0.500
Pain medication	−0.041	0.005	5	0.192
Healthcare expenditure (>p50)	−0.001	0.004	12	0.462
Panel C: Part-time teachers				
BAPO take-up	0.344	0.001	1	0.038
Retention	−0.027	0.003	3	0.115
Total teaching supplied	−0.067	0.003	1	0.038
Antidepressants	0.012	0.001	8	0.308
Pain medication	−0.066	0.001	3	0.115
Healthcare expenditure (>p50)	0.004	0.007	14	0.538

Notes: The table re-estimates equation (1) at 25 placebo cutoffs, monthly birth-date thresholds from May 1963 to May 1965 at which eligibility does not change; all placebo cohorts are deferred to age 57, and the nine-month gap after the true cutoff ensures that no placebo window overlaps the true discontinuity. True estimate is the RD estimate at the August 1, 1962, cutoff and Placebo mean the mean of the 25 placebo estimates. Rank is the position of the true estimate among the 26 estimates, counted from the largest for positive estimates and from the smallest for negative ones; p is the corresponding one-sided randomization-inference p-value, rank/26 (Ganong and Jäger, 2018), whose smallest attainable value is $1/26 \approx 0.038$. Outcomes are pooled over 2015–2018 (2016–2018 for total teaching supplied) with a 240-day bandwidth, a local linear specification, and a uniform kernel. Health outcomes are not conditional on remaining in teaching. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below.

Table A.13: Robustness: baseline covariates included

	RD estimate		
	All	Full-time	Part-time
	(1)	(2)	(3)
Panel A: Extensive margin			
BAPO take-up	0.406***	0.468***	0.358***
	(0.030)	(0.049)	(0.038)
Retention	0.015	0.059**	−0.017
	(0.020)	(0.028)	(0.027)
Observations	8,747	3,740	5,007
Panel B: Intensive margin			
Teaching supplied (FTE)	−0.004	0.040	−0.039
	(0.020)	(0.032)	(0.024)
Observations	6,524	2,787	3,737
Panel C: Health			
Antidepressants	0.020	0.019	0.021
	(0.017)	(0.023)	(0.023)
Pain medication	−0.052**	−0.034	−0.066**
	(0.021)	(0.033)	(0.026)
Healthcare expenditure (>p50)	0.006	0.011	0.004
	(0.030)	(0.046)	(0.041)
Observations	8,747	3,740	5,007

Notes: The table reports RD estimates of equation (1) including the full vector of baseline covariates: gender, immigrant background, contractual FTE, and antidepressant and pain-medication use, all measured in 2013. Outcomes are defined as in Tables 2–4 and pooled over 2015–2018 (2016–2018 for Teaching supplied (FTE) in panel B). Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.14: Local randomization estimates of the effect of eligibility on BAPO take-up and retention

	Window (days from cutoff)		
	±30	±60	±90
Panel A: All teachers			
BAPO take-up	0.423***	0.409***	0.396***
	(0.000)	(0.000)	(0.000)
Retention	0.040	0.016	0.019
	(0.205)	(0.433)	(0.267)
Obs. (left/right)	137 / 144	276 / 282	406 / 436
Panel B: Full-time teachers			
BAPO take-up	0.482***	0.423***	0.460***
	(0.000)	(0.000)	(0.000)
Retention	0.061	0.041	0.055**
	(0.177)	(0.176)	(0.016)
Obs. (left/right)	65 / 60	131 / 121	186 / 185
Panel C: Part-time teachers			
BAPO take-up	0.385***	0.400***	0.350***
	(0.000)	(0.000)	(0.000)
Retention	0.027	−0.002	−0.008
	(0.574)	(0.978)	(0.743)
Obs. (left/right)	72 / 84	145 / 161	220 / 251

Notes: The table reports local randomization estimates of the effect of eligibility within symmetric windows of ±30, ±60, and ±90 days around the August 1, 1962, cutoff, following Cattaneo et al. (2016). Each cell reports the difference in mean outcomes between teachers born before and on or after the cutoff, pooled over 2015–2018; finite-sample randomization p-values, based on 10,000 permutations of eligibility within the window, are in parentheses. Obs. (left/right) is the number of teachers on each side of the cutoff within the window. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.15: Robustness: MSE-optimal bandwidths with robust bias-corrected inference

	RD estimate	Robust std. err.	Robust p-value	h_ℓ	h_r	N_ℓ	N_r
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: All							
BAPO take-up	0.391***	(0.039)	<0.001	253.5	253.5	4,562	4,677
Retention	0.023	(0.023)	0.213	306.5	306.5	5,463	5,787
Teaching supplied (FTE)	−0.024	(0.026)	0.417	345.6	345.6	4,583	4,907
Antidepressants	0.006	(0.022)	0.671	327.8	327.8	5,848	6,247
Pain medication	−0.058**	(0.029)	0.019	209.4	209.4	3,731	3,852
Healthcare expenditure (>p50)	−0.006	(0.037)	0.498	309.3	309.3	5,503	5,838
Panel B: Full-time							
BAPO take-up	0.449***	(0.060)	<0.001	261.4	261.4	1,971	2,153
Retention	0.072**	(0.035)	0.017	260.6	260.6	1,967	2,137
Teaching supplied (FTE)	0.051*	(0.039)	0.090	244.9	244.9	1,385	1,462
Antidepressants	0.009	(0.026)	0.706	395.8	395.8	2,905	3,439
Pain medication	−0.035	(0.038)	0.290	355.2	355.2	2,656	3,052
Healthcare expenditure (>p50)	−0.017	(0.054)	0.756	306.8	306.8	2,271	2,615
Panel C: Part-time							
BAPO take-up	0.352***	(0.046)	<0.001	288.4	288.4	2,991	2,930
Retention	−0.013	(0.027)	0.672	437.9	437.9	4,414	4,493
Teaching supplied (FTE)	−0.046	(0.031)	0.124	359.0	359.0	2,738	2,776
Antidepressants	0.002	(0.031)	0.825	333.6	333.6	3,464	3,455
Pain medication	−0.088**	(0.039)	0.011	176.5	176.5	1,776	1,840
Healthcare expenditure (>p50)	0.002	(0.047)	0.992	319.4	319.4	3,355	3,296

Notes: The table reports RD estimates at the MSE-optimal bandwidth selected separately by outcome and sample following Calonico et al. (2014), pooled over 2015–2018 (2016–2018 for Teaching supplied (FTE)). Column 1 reports the conventional local linear estimate; columns 2 and 3 the robust bias-corrected standard error and p-value, on which significance levels are based. Columns 4 and 5 report the bandwidths in days to the left and right of the cutoff, which coincide under the common-bandwidth procedure; columns 6 and 7 the corresponding numbers of observations. Estimates use a triangular kernel, with standard errors clustered at the teacher level. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.16: Robustness: second-order polynomial in the running variable

	RD estimate		
	All	Full-time	Part-time
	(1)	(2)	(3)
Panel A: Extensive margin			
BAPO take-up	0.390***	0.406***	0.386***
	(0.046)	(0.074)	(0.058)
Retention	0.052*	0.083*	0.030
	(0.031)	(0.043)	(0.043)
Observations	8,747	3,740	5,007
Panel B: Intensive margin			
Teaching supplied (FTE)	−0.012	0.059	−0.033
	(0.036)	(0.047)	(0.044)
Observations	6,524	2,787	3,737
Panel C: Health			
Antidepressants	−0.003	−0.005	−0.006
	(0.030)	(0.041)	(0.044)
Pain medication	−0.064**	−0.032	−0.081**
	(0.032)	(0.053)	(0.039)
Healthcare expenditure (>p50)	−0.047	−0.048	−0.050
	(0.047)	(0.070)	(0.062)
Observations	8,747	3,740	5,007

Notes: The table reports RD estimates of equation (1) allowing a second-order polynomial in the running variable on each side of the cutoff. Outcomes are defined as in Tables 2–4 and pooled over 2015–2018 (2016–2018 for Teaching supplied (FTE) in panel B); control means are reported in the main tables. Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days around the August 1, 1962, cutoff with a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.17: Robustness: donut regression discontinuity estimates

	RD estimate		
	All	Full-time	Part-time
	(1)	(2)	(3)
Panel A: Extensive margin			
BAPO take-up	0.373***	0.463***	0.309***
	(0.040)	(0.063)	(0.050)
Retention	−0.009	0.058	−0.059*
	(0.025)	(0.036)	(0.035)
Observations	7,632	3,241	4,391
Panel B: Intensive margin			
Teaching supplied (FTE)	−0.027	0.047	−0.068*
	(0.030)	(0.042)	(0.035)
Observations	5,689	2,413	3,276
Panel C: Health			
Antidepressants	−0.004	−0.009	−0.000
	(0.024)	(0.033)	(0.034)
Pain medication	−0.034	−0.012	−0.050
	(0.029)	(0.047)	(0.036)
Healthcare expenditure (>p50)	0.039	0.041	0.035
	(0.041)	(0.064)	(0.054)
Observations	7,632	3,241	4,391

Notes: The table reports RD estimates of equation (1) excluding teachers born within one month of the August 1, 1962, cutoff. Outcomes are defined as in Tables 2–4 and pooled over 2015–2018 (2016–2018 for Teaching supplied (FTE) in panel B). Full-time is defined as ≥ 0.95 FTE in 2013; part-time, below. The bandwidth is 240 days, with a local linear specification and a uniform kernel. Standard errors, in parentheses, are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.18: Placebo difference-in-differences: testing for a control-cohort response

	Retention	Antidepressants	Pain medication
	(1)	(2)	(3)
Panel A: All			
Cohort 1 × Post	−0.019*	0.002	−0.003
	(0.010)	(0.007)	(0.011)
Cohort 2 × Post	−0.008	0.002	−0.010
	(0.010)	(0.007)	(0.012)
Observations	25,019	25,019	25,019
Panel B: Full-time			
Cohort 1 × Post	−0.010	−0.001	0.025
	(0.015)	(0.008)	(0.017)
Cohort 2 × Post	0.002	0.000	−0.026
	(0.014)	(0.009)	(0.018)
Observations	10,799	10,799	10,799
Panel C: Part-time			
Cohort 1 × Post	−0.026*	0.004	−0.024
	(0.014)	(0.011)	(0.015)
Cohort 2 × Post	−0.016	0.004	0.003
	(0.014)	(0.010)	(0.016)
Observations	14,220	14,220	14,220

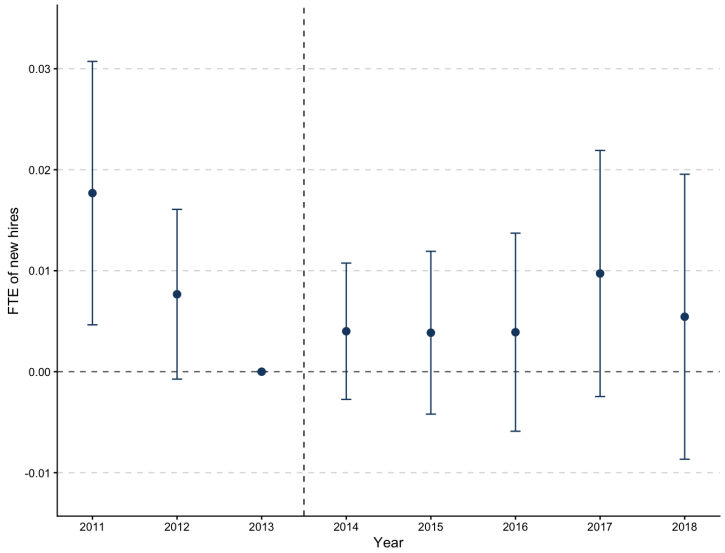
Notes: Panel A pools all teachers; Panel B restricts the sample to full-time teachers ($\geq$ 0.95 FTE in 2013); Panel C to part-time teachers. Each column reports a difference-in-differences across three cohorts excluded from the transitional BAPO arrangement, estimated over 2011–2018. Cohort 1 is the RD control cohort (born August 1962–March 1963), which missed eligibility most narrowly; Cohort 2 is born one year later (August 1963–March 1964); the omitted reference cohort is born two years later (August 1964–March 1965). Post equals one for 2015–2018. All specifications include cohort and year fixed effects. If narrowly missing the cutoff generates a disappointment response, it should be strongest for Cohort 1 and attenuate with distance from the cutoff. Coefficients not statistically distinguishable from zero, and showing no gradient across cohorts, indicate no such response. All outcomes are defined as in Tables 2 and 4. Standard errors in parentheses are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Table A.19: Difference-in-differences estimates of school exposure on incumbent teachers' working hours, by age group

	Young (< 35) (1)	Middle-aged (35–49) (2)	Older (≥ 50) (3)
Exposure (10 pp) × Post	0.0175***	0.0017	0.0002
	(0.0047)	(0.0024)	(0.0024)
Observations	163,124	243,079	199,663
Teachers	22,707	32,379	30,778

Notes: Each column reports the coefficient on $\pi_s \times \text{Post}_t$ from a teacher fixed-effects regression of actual working FTE (set to zero for teachers no longer employed as teaching staff) on school exposure interacted with a post-2014 indicator, controlling for the 2013 FTE-weighted shares of the deferred, always-eligible, and younger cohorts and school size, each interacted with Post. Coefficients and standard errors are scaled to a 10 percentage point increase in the exposure share. Age is defined in 2013. Standard errors in parentheses are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

Figure A.11: Event-study estimates of school exposure on new hiring



Notes: The figure plots event-study coefficients from the school-level analogue of equation (3), with school and year fixed effects and 2013 as the omitted reference year. Coefficients are scaled to a 10 percentage point increase in the exposure share π_s. The outcome is the FTE of teachers employed at the school in year t who were not employed there in year $t-1$, relative to 2013 capacity, constructed from the teacher register. The sample is restricted to schools with at least 15 teachers in 2013. Bars denote 95 percent confidence intervals based on standard errors clustered at the school level.

B Sample Construction

My primary data source is the administrative registers of public-sector education employees in the Netherlands, provided by DUO. Spanning 2006 to 2024, these registers record every individual employment relationship as observed in October of each year. I begin from all individuals employed in secondary education in 2013 and born within my estimation bandwidth (December 1961 to March 1963), yielding 3,854 job records for 3,787 individuals. Jobs are classified as teaching (2,407 records, 62.5 percent), educational support and administration (1,307 records, 33.9 percent), management (138 records, 3.6 percent), or other (2 records, 0.1 percent).

The main analysis focuses on teachers, defined as all individuals holding at least one teaching job in 2013. This selection retains 2,412 job records for 2,362 individuals: the 2,407 teaching jobs plus five non-teaching jobs held by these same teachers. The 50 additional records (2.1 percent) belong to teachers holding multiple jobs, either within the same school or across schools. I aggregate them to a single record per teacher, summing FTE and salary across jobs and assigning each teacher to the school accounting for the largest number of hours, and I apply the same rule in all subsequent years. To validate the school-reported records, I link them to the employer–employee register (SPOLISBUS), drawn from firm tax records. Seven individuals appear in the teacher register but not in the tax data. I treat these as school reporting errors, most plausibly teachers no longer holding a formal contract with the school, and drop them, leaving 2,355 teachers. The registers also capture individuals whose teaching contract sits alongside larger employment (or self-employment) elsewhere. Because the analysis concerns teaching careers, I restrict the sample to teachers for whom teaching was the main source of income in 2013. This drops 77 individuals, leaving the 2,278 teachers of the main analysis.

I then construct a balanced panel over 2015–18 for these 2,278 teachers, for a total of 9,112 person-year observations. Over this period, 43 individuals cease to be Dutch residents, chiefly through death. Because the health data cover only residents insured under the mandatory basic health-insurance scheme, I drop these individuals entirely, including the years in which they were still resident, so that labor-market and health outcomes are estimated on an identical sample. This leaves 8,940 observations on 2,235 teachers. An RD estimate confirms that eligibility does not affect the probability of remaining resident (Table A.1), so the restriction is unlikely to introduce differential selection at the cutoff.

Finally, 193 of the remaining person-years (2.2 percent) appear in the employer–employee data under the education sector but not in the teacher register. This typically

reflects delayed school reporting to DUO, which the October snapshot misses: an individual may appear continuously in the employer–employee data under the same encrypted firm identifier while being absent from the teacher register in a single year between 2010 and 2018. Because retaining these person-years would require assuming continued employment as a teacher, the main specification drops them, yielding a final, unbalanced estimation panel of 8,747 observations. As a robustness check, I instead impute them as teaching employment, taking the job function from the teacher register and hours and salary from the employer–employee data; the results are unchanged (Table A.2).

C The Netherlands Working Conditions Survey (NEA)

The NEA (*Netherlands Working Conditions Survey*) is an annual survey of employees aged 15–74 working in the Netherlands, excluding the self-employed, conducted by TNO and Statistics Netherlands in cooperation with the Ministry of Social Affairs and Employment. It covers working conditions, job content, and employment terms. The survey is a repeated cross-section, and from 2017 onward the sample expands substantially, increasing the number of respondents in occupations such as teaching. Throughout, I restrict attention to secondary-school teachers aged 48–57 and born in or after August 1962, so that all face the post-reform regime without transitional BAPO access, and I pool all available waves (2014–2024). [31]

I use these data to compare job strain and the willingness to reduce working hours. The comparison between full-time and part-time teachers (Table C.1) uses all NEA respondents identified as secondary-school teachers in this age range; linking them to the analysis sample would leave too little overlap. The comparison between teachers and support staff (Table C.2) instead links NEA respondents through unique personal identifiers to the DUO personnel register, restricting to personnel employed in secondary education. Teaching and administrative and educational-support status is taken directly from the register, since the NEA itself cannot distinguish the two.

Variables are defined as follows. *Willing to work longer with fewer hours* equals one if the respondent reports that working fewer hours or fewer days per week is a circumstance under which they would be willing to continue working to a higher age. The item belongs to a multiple-response battery on conditions for working longer and is not fielded in 2014, 2016, 2018, or 2020, which accounts for its smaller sample size. *High work pressure* equals

[31] Figure A.1 is the exception: it compares all secondary-school teachers with the Dutch workforce by age group over 50–65.

one if the respondent scores at least 2.5 on a three-item scale (working very fast, doing a great deal of work, working extra hard; 1 = never to 4 = always). *Physically demanding work* is a derived indicator based on exerting force, vibrations, awkward postures, and repetitive movements. *Emotionally demanding work* equals one if the respondent scores at least 2.5 on a three-item scale (emotionally difficult situations, emotionally demanding work, emotional involvement; same scale).

Table C.1: Workload and hours preferences of full- and part-time teachers (NEA)

	Part-time		Full-time		
	Mean	N	Mean	N	p-value
	(1)	(2)	(3)	(4)	(5)
Willing to work longer with fewer hours	0.657	1,459	0.738	690	0.000
High work pressure	0.484	2,209	0.520	981	0.061
Physically demanding work	0.101	2,198	0.132	980	0.012
Emotionally demanding work	0.263	1,887	0.282	794	0.304

Notes: The table reports means of NEA survey indicators for secondary-school teachers aged 48–57 and born in or after August 1962, pooled over 2014–2024. Variables are defined in the text. Column (5) reports the p-value of a two-sample test of equal means. Full-time status is defined from contractual weekly hours, approximating the ≥ 0.95 FTE definition used in the administrative analysis. The number of observations varies because of item non-response and because *willing to work longer with fewer hours* is not fielded in 2014, 2016, 2018, and 2020.

Table C.2: Working conditions and characteristics of lowest-scale teachers and support staff (NEA linked to the personnel register)

	Teachers, scale LB		Support staff		
	Mean	N	Mean	N	p-value
	(1)	(2)	(3)	(4)	(5)
Panel A: Working conditions (NEA)					
Willing to work longer with fewer hours	0.667	102	0.485	241	0.002
High work pressure	0.512	242	0.374	487	<0.001
Physically demanding work	0.138	240	0.243	489	<0.001
Emotionally demanding work	0.326	239	0.096	470	<0.001
Panel B: Characteristics in 2013					
Contractual working hours (FTE)	0.795	243	0.800	489	0.764
Monthly salary (euros)	2,882	243	2,403	489	<0.001
Female	0.584	243	0.603	489	0.624
Immigrant background	0.210	243	0.104	489	<0.001
Antidepressants	0.070	243	0.061	489	0.655
Pain medication	0.218	243	0.198	489	0.534

Notes: The table reports means of NEA survey indicators (Panel A) and of 2013 register and health-insurance characteristics (Panel B) for NEA respondents linked to the DUO personnel register, restricted to secondary-education personnel aged 48–57 and born in or after August 1962, pooled over 2014–2024. Columns (1)–(2) are restricted to teachers in salary scale LB; staff category is taken from the register. Column (5) reports the p-value of a two-sample test of equal means. Variables are defined as in Table C.1 and Table 1. The number of observations varies because of item non-response and because *willing to work longer with fewer hours* is not fielded in 2014, 2016, 2018, and 2020.

D Conditional Working Hours

Table D.1 and Figure D.1 report effects on actual and contractual working hours, conditional on remaining a teacher. Eligibility reduces actual hours by 0.023 FTE among full-time teachers and by 0.048 FTE among part-time teachers. Because compliers in both groups use the same share of their entitlement, the smaller reduction in actual hours among full-time teachers must reflect a response in contractual hours. Consistent with this interpretation, contractual hours increase by 0.023 FTE at the cutoff among full-time teachers.

This increase cannot reflect eligible teachers expanding their contracts: the collective agreement caps contracts at 1.0 FTE while BAPO leave is being taken (VO-raad, 2014). Nor can it be explained by changes in the composition of the surviving teacher pool. The retention response can account for at most 0.014 FTE of the increase, and only under the extreme assumption that every marginal stayer holds the largest permissible contract and takes no leave.[32] The remaining increase therefore reflects ineligible full-time teachers reducing their contracts at their own expense.

This pattern suggests that some full-time teachers are willing to reduce their workload even when doing so is costly, while the teachers deterred by that cost are precisely those whom eligibility induces to remain. The magnitudes are consistent with this interpretation. Thus, part of the subsidy appears to finance reductions in working time that some eligible teachers would otherwise have purchased themselves.

Part-time teachers show no comparable contractual response. Their contractual-hours estimate is negative and imprecisely estimated, while actual hours fall by 0.048 FTE. This pattern is consistent with funded leave passing through to Teaching supplied (FTE) for part-time teachers, rather than subsidizing reductions in contractual hours that they would otherwise have financed themselves.

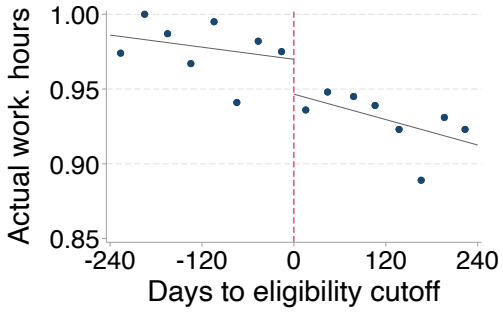
[32]Marginal stayers make up a share $s = \Delta/S_1 = 0.061/0.961 \approx 0.063$ of eligible full-time survivors, where $\Delta = 0.061$ is the retention effect and $S_1 = 0.900+0.061 = 0.961$ the implied retention rate of eligible teachers. Their maximum contribution to mean contractual hours is therefore $s\,(\bar{C}_{\text{marginal}} - \bar{C}_{\text{infra}}) \leq 0.063 \times (1.2 - 0.983) \approx 0.014$ FTE, where $\bar{C}_{\text{infra}} = 0.983$ is the control mean and 1.2 FTE is the maximum contract for a teacher not taking leave. The same share bounds any compositional effect on the leave-share estimates in Section 5 at $0.063 \times 0.102 \approx 0.006$, too small to affect the conclusion that takers exhaust the entitlement.

Table D.1: Net effect of eligibility on working hours, by employment group

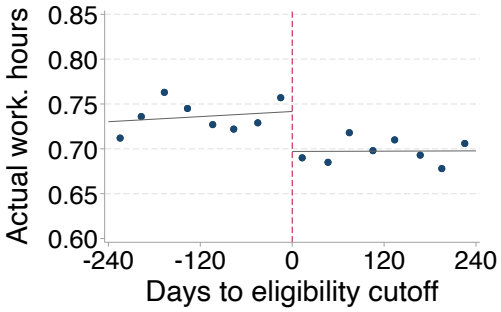
	Contractual working hours (1)	Actual working hours (2)
Panel A: All		
RD estimate	−0.009	−0.043***
	(0.015)	(0.015)
Control mean	0.855	0.855
Observations	5,921	5,921
Panel B: Full-time		
RD estimate	0.023***	−0.023**
	(0.009)	(0.010)
Control mean	0.983	0.983
Observations	2,595	2,595
Panel C: Part-time		
RD estimate	−0.024	−0.048***
	(0.018)	(0.017)
Control mean	0.746	0.746
Observations	3,326	3,326

Notes: The table reports RD coefficients estimated from equation (1). The sample consists of secondary-school teachers employed in 2013 and is restricted to teachers observed in the teacher register in 2016–2018, the period for which leave hours are recorded. *Contractual working hours* are the FTE specified in the employment contract. *Actual working hours* are contractual hours net of subsidized leave, expressed in FTE. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below. The bandwidth is 240 days around the 1 August 1962 cutoff, with a local linear specification and a uniform kernel. Standard errors in parentheses are clustered at the teacher level. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$.

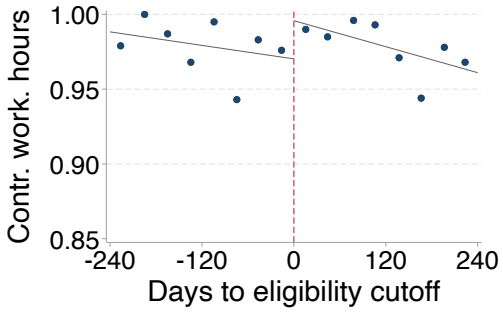
Figure D.1: Effect of eligibility on working hours, by contract type



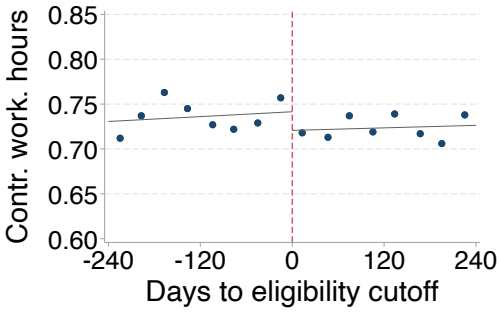
(a) Full-time: actual working hours



(b) Part-time: actual working hours



(c) Full-time: contractual working hours



(d) Part-time: contractual working hours

Notes: Panels (a) and (b) plot actual working hours, defined as contractual FTE net of subsidized leave, and panels (c) and (d) contractual working hours in FTE, against date of birth relative to the 1 August 1962 cutoff, pooled over 2016–2018, the period for which leave hours are recorded. All outcomes are conditional on being observed in the teacher register. Bins are evenly spaced over 30 days with separate linear fits on each side; the bandwidth is 240 days with a uniform kernel. Full-time is defined as ≥ 0.95 FTE in 2013; part-time below.

E Cost-Effectiveness: Details

E.1 The cost of a vacancy

When a teacher leaves, her school must cover her lessons until the position is filled. In October 2024, 31 percent of teacher shortage FTE in secondary education were open vacancies, 51 percent were covered internally, and 17 percent by agency or self-employed staff (Ministerie van Onderwijs, Cultuur en Wetenschap, 2024, Table 11). Table E.1 values each mode relative to keeping the senior teacher, in money and in student achievement.

In money terms, internal cover is cheaper than the senior: I assume that colleagues covering the lessons earn 85 percent of senior pay. Agency cover is more expensive: a

seconded teacher costs €113 per hour, against €58 for a directly employed teacher in the same scale and step (Algemene Rekenkamer, 2024). Open vacancies are treated as covered internally, since their lessons are patched by colleagues while the position remains unfilled. Weighting each mode by its share, a vacancy-year costs about 5 percent of a salary more than keeping the senior.

In terms of achievement, the main cost of a vacancy is that lessons are often taught by teachers not qualified in the subject. Three quarters of internal cover, as well as the open vacancies patched internally, are taught out of field, so 77 percent of shortage FTE are taught by a teacher without a qualification in the subject. Students of such teachers score 0.05–0.11 standard deviations lower (Clotfelter et al., 2010). Assuming that qualified colleagues and qualified agency staff are as effective as the senior, a vacancy-year lowers achievement by about 0.06 standard deviations.

Table E.1: Cost of covering a vacancy-year, by coverage mode

Coverage mode	Weight	Money (salaries)	Quality loss (SD)	Quality loss, midpoint
Open vacancy, patched out-of-field	0.31	−0.15	0.05–0.11	0.08
Qualified colleague, extra hours	0.12	−0.15	0	0
Out-of-field or unqualified colleague	0.39	−0.15	0.05–0.11	0.08
Agency or self-employed, qualified	0.11	+0.95	0	0
Agency or self-employed, out-of-field	0.07	+0.95	0.05–0.11	0.08
Weighted average		+0.05		0.06

Notes: Money is the cost of cover minus the senior's salary, in units of €51,300: a seconded teacher costs €113 per hour against €58 for a directly employed teacher in the same scale and step (Algemene Rekenkamer, 2024); colleagues covering internally are assumed to earn 0.85 of senior pay. Weights are shares of total teacher shortage FTE in October 2024 (Ministerie van Onderwijs, Cultuur en Wetenschap, 2024, Table 11); the remaining share (merged classes) is omitted, and weights do not sum exactly because of rounding. Open vacancies are assigned to out-of-field cover, since their lessons are patched internally while they remain unfilled. Quality loss is in standard deviations of student test scores per year of exposure, relative to an experienced teacher certified in the subject. Out-of-field: students of a teacher not certified in the subject score 0.07 SD lower on average (0.05 if certified in a related subject) and 0.10–0.11 lower in mathematics and English (Clotfelter et al., 2010, Tables 4–5). Qualified colleagues and qualified agency staff are assumed to be as effective as the senior. Weighted averages use midpoints.

E.2 Valuing achievement

I value achievement with Chetty et al. (2014b): a one-year gain of 0.10 student standard deviations is worth about $7,000 per student in lifetime earnings, the present value at age 12 of a 1.3 percent earnings gain (2010 dollars, 5 percent discount rate) for one standard deviation of teacher value added, which equals about 0.10 student standard deviations (Chetty et al., 2014a). A full-time teacher teaches 22 periods per week; at about three periods per class and 20–25 students per class, she teaches roughly 140–175 students. Across 150 students, 0.01 standard deviations is therefore worth about €140,000 in 2024 prices. The euro values scale proportionally with the number of students; the break-even thresholds below do not. The estimates come from US grades 4–8 and project earnings beyond age 28, so their magnitude may not transfer to Dutch secondary schools.

The avoided vacancy is worth 6× €140,000 ≈ €840,000 per retained teacher-year. Leave hours are covered mainly by incumbents under 35 (Section 7). Teachers in their first five years are 0.02–0.06 standard deviations less effective than experienced teachers (Rivkin et al., 2005; Clotfelter et al., 2010). Assuming a third of the covering teachers are in their first five years, covered hours lose about 0.013 standard deviations per FTE, or about €182,000.

E.3 Full calculation

Table E.2 reports all components per 100 eligible teachers (43 full-time and 57 part-time) per year. Part-time retention is set to zero, since its estimate is insignificant.

Table E.2: Costs and benefits per 100 eligible teachers per year

	Full-time (43)	Part-time (57)	Total (100)
Panel A: Money			
Subsidy on leave hours	−65,000	−38,000	−103,000
Leave covered by colleagues (FTE)	1.9	1.2	3.1
Wage saving on covered hours	+15,000	+9,000	+24,000
Retained teacher-years	2.6	0	2.6
Avoided vacancy cost	+7,000	0	+7,000
Net, money	−43,000	−29,000	−72,000
Panel B: Achievement			
Avoided achievement loss	+2,180,000	0	+2,180,000
Achievement loss on covered hours	−350,000	−220,000	−570,000
Net, including achievement	+1,790,000	−250,000	+1,540,000

Notes: Euros per year, 2024 prices. Leave FTE equals the leave share among stayers times the retention rate of eligible teachers times mean contract size (1.0 and 0.70). Retained teacher-years equal 43×0.061. Covering colleagues earn 0.85 of senior pay (salary per FTE €51,300 for full-time and €48,900 for part-time teachers) and are 0.013 SD less effective per FTE. Each retained teacher-year avoids a vacancy-year costing 0.05 of a salary and 0.06 SD (Appendix E.1). Achievement is valued at €140,000 per 0.01 SD per full-time teacher-year (Appendix E.2).