SAPIENZA
UNIVERSITÀ DI ROMA

Sapienza University of Rome

Department of Economics and Law
PhD in Economics

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Essays in Applied Labor Economics: Causal Inference on Gender, Time Allocation, and Adult Children Fertility at the Retirement Threshold

Thesis Supervisors
Prof.ssa Emanuela Ghignoni
Prof. Marco Ventura

Candidate
Marianna Nitti
1556797

Academic Year 2024-2025 (XXXVII cycle)

To Ale

Contents

Summary

The ageing of the population and the sustainability of welfare systems pose crucial challenges for European economies, making retirement a pivotal event at both the individual and policy levels. The transition from work to pension not only redefines an individual's life but also triggers profound transformations in family dynamics, with significant gender and intergenerational implications.

The challenge this dissertation seeks to tackle is twofold. First, it aims to estimate the causal effect of retirement on two key dimensions of family behaviour: time allocation and fertility choices. Specifically, it analyses how men and women reorganise their time between domestic work and leisure after retirement, and how the newly found time availability of grandparents influences the reproductive decisions of their children. To credibly isolate these causal effects, the dissertation adopts a Regression Discontinuity Design (RDD), using statutory pension eligibility ages as a source of exogenous variation. Second, the dissertation contributes to the advancement of this methodology by developing tools that integrate Machine Learning techniques to enhance the precision and robustness of estimates in the presence of complex data.

The manuscript is structured as a sequence of three chapters, all connected by the adoption of the RDD framework to analyse socio-economic phenomena. Chapter 1 applies a fuzzy RDD to assess the causal effect of retirement on time allocation between men and women in Italy, examining the gender implications for healthy ageing. Chapter 2 presents a purely methodological contribution, developing a new Stata command, rdlasso, which implements a Lasso-based procedure for selecting high-dimensional covariates in both sharp and fuzzy RDD settings. Finally, Chapter 3 extends the analysis to a pan-European context, using the RDD framework to study the impact of grandparental retirement on the fertility of their adult children, testing the heterogeneity of this causal link across different welfare systems.

In chronological terms, each chapter of this doctoral thesis represents a modest advance over the previous. While Chapter 1 rigorously applies the RDD methodology to a critical social policy question, Chapter 2 directly addresses a methodological challenge by creating a practical tool to enhance the technique itself. Chapter 3, in turn, synthesises the previous two approaches by applying an RDD methodology, potentially enhanced by the techniques discussed in the second chapter, to a complex, large-scale causal analysis, thereby demonstrating the flexibility and robustness of the proposed methodological framework.

Chapter one

Gender differences in time allocation after retirement: a regression discontinuity approach

Emanuela Ghignoni * Marianna Nitti * Marco Ventura*

**Sapienza, University of Rome*

Abstract

The rise in women's participation in the labour market has resulted in only a minor shift in the division of household labour, with (working) women predominantly bearing the responsibility for housework. This paper focuses on whether we should expect a rebalancing or a reinforcement of gender disparities in time allocation after retirement.

We employ a fuzzy regression discontinuity design (RDD) to Italian data, using retirement eligibility as an exogenous instrument for retirement status, to identify the causal effect of retirement on time allocation for men and women. Our results show a strong first stage, indicating substantial compliance with retirement rules for both men and women. For men, retirement mainly reduces paid work and total work, with leisure emerging as an important margin of time reallocation. For women, by contrast, retirement is associated with a reduction in unpaid work and an increase in leisure time, pointing to a rebalancing of post-retirement time use across genders. This represents a novel and particularly intriguing finding for women, with potentially important policy implications.

Keywords: Time allocation, Gender, Retirement, Healthy ageing, Fuzzy Regression Discontinuity Design

JEL Codes: H51, J22, J26, J16, J14

1 Introduction and relevant literature

The Italian population has been rapidly ageing for several decades, a trend that shows no signs of reversing. This demographic transition has led to a steady increase in the number of pension recipients, the majority of whom are women, and has placed substantial pressure on public finances. In response, successive pension reforms have progressively increased statutory retirement ages and tightened contribution requirements in an attempt to ensure the sustainability of the pension system.

At the same time, Italy is characterised by persistent gender disparities in labour market participation and in the allocation of unpaid work. According to the Italian National Institute of Statistics (Istat, 2021), the female employment rate is about 20 percentage points lower than that of men (55.7% versus 75.8%), with even wider gaps in Southern regions. Moreover, almost one third of employed women work part-time.

According to European Commission (2021), despite the unadjusted gender pay gap in Italy being 5% (below the European average which is 12.7%), women continue to be at greater risk of poverty, carry out more than double of unpaid work compared to men and experience less leisure than men. At the same time, women (particularly the highly educated) have slightly reduced unpaid work and increased leisure.

These inequalities in time allocation are well documented during working life, but much less is known about how they evolve after retirement. Understanding how men and women reallocate their time after leaving the labour market is important for several reasons. First, retirement represents a major life transition that relaxes the time constraints previously imposed by paid employment. For this reason, individuals may reallocate time away from market work toward unpaid work, leisure, or other activities. Second, the way this newly available time is reallocated may either reduce or reinforce pre-existing gender disparities in time use. For this reason, retirement provides a unique setting to examine whether gender inequalities persist even when labour market constraints are removed. Finally, despite its importance, this question remains largely unanswered, particularly in a causal framework.

A large literature has documented that working women, even when employed full-time, remain disproportionately responsible for domestic production and care work throughout their careers. For Italy, Addabbo et al. (2012) analyse time allocation among individuals aged 18 to 60 and find that the presence of children reduces women's paid work while increasing unpaid work for both partners. Barigozzi et al. (2023), using a long time span, show that while gender gaps in market and domestic work among young couples with children have narrowed over time, gaps in childcare and leisure have remained largely unchanged. However, very few studies have investigated the changes in unpaid work and leisure that occur among retirees of both genders.

Stancanelli and Soest (2012), using French data, exploit a sharp age-based cutoff at the minimum age for early retirement (60) to identify the effect of retirement on market work, home production, and caring activities. They find a substantial increase in home production after retirement for both men and women, with larger effects for men. Similarly, Bonsang and Van Soest (2020) study Germany and rely on age-based eligibility thresholds to identify the impact of retirement on household production within couples, documenting significant increases in home production following retirement.

For Italy, Ciani (2016) estimates the effect of retirement on time devoted to housework. Eligibility is reconstructed by combining age, years of contributions, and work history, and the running variable is defined as the (year-level) distance to eligibility. The analysis shows that women substantially increase unpaid work after retirement, while effects for married men are close to zero.

Importantly, most of these studies either focus on a limited set of time-use outcomes or rely on age-only cutoffs measured in years, which may introduce imprecision in the definition of eligibility.

This paper contributes to the literature by providing causal evidence on gender differences in time reallocation after retirement in Italy. We focus on how retirement affects paid work, unpaid work, leisure, and total work time for men and women.

To identify causal effects, we exploit the discontinuities in retirement probabilities generated by statutory age and contribution requirements embedded in the Italian pension system. We implement a fuzzy regression discontinuity design, using retirement eligibility as an exogenous instrument for retirement status.

Compared with previous contributions that rely on age-only cutoffs with age measured in years as the running variable, we construct a month-level running variable that integrates statutory age and contributory requirements across pension reforms. Coupled with monthly age information, this approach reduces misclassification of eligibility status and increases the precision of the estimates. In addition, we exploit recent Italian time-use data (SHARE Wave 8, SHARE-ERIC (2022b)), where economic-status information and complete 24-hour time-use coexist in a single dataset, allowing consistent outcome definitions and avoiding imputed or proxied measures. Finally, on the methodological side, we augment standard RDD diagnostics with external-validity tests, which enhance the economic interpretation of the econometric results, such as the stability of the compliers population.

Our findings indicate that retirement leads to a reduction in total work time for both men and women. Nevertheless, the reduction in total work for men is predominantly attributable to a decline in paid employment, whereas for women it reflects a decrease in unpaid work and an increase in leisure time. This represents a novel finding for women. Previous studies, in fact, typically document a significant increase in unpaid work for retired female workers, rather than gains in leisure time.

Another noteworthy finding is the large proportion of compliers at the cutoff, i.e. individuals who decide to retire as soon as they become eligible. Our estimates are higher than those reported by Battistin et al. (2009) for the period 1993–2004 and highlight the strong effect that successive pension reforms have had over the last decades.[1] Overall, these results reflect the cumulative impact of past pension reforms.

Both sets of findings are robust across specifications and to a battery of internal and external validity tests. From a policy perspective, the results indicate that statutory eligibility rules strongly shape retirement behaviour for both men and women, with high take-up rates at the

[1] Our estimates are closer to those reported by Marini (2024), who documents a take-up rate of about 70% over the period 2010–2016.

threshold.

In principle, even marginal changes in eligibility conditions may translate into sizable shifts in retirement behaviour, with potentially important implications for pension expenditures and the distribution of pension entitlements.

The remainder of the paper is structured as follows. Section 2 explains the institutional setup of the Italian context, Section 2 presents the data, and Section 3 describes the identification strategy, empirically tested in Section 3.2. Results are discussed in Section 4 and the external validity tests are reported in Section 6. Section 5 concludes and provides some policy recommendations.

2 Institutional framework

2.1 The pension system in Italy

Broadly speaking, there are two types of pensions: the *old-age pension* and the *early-retirement pension*. The latter definition is quite recent, as this type of pension was previously called *seniority pension*.

In the case of old-age pension, the individuals receive the pension at the end of their working life when the necessary requirements are met. The second type of pension consists of an early exit from the labour market reserved for individuals who meet certain requirements in terms of amount of years of contribution.

Pensions also differ on the basis of gender and the type of work performed during working life (private employees, public employees or the self-employed).

As a result of the large impact that the pension system has had on public finances, numerous legislative measures have been introduced over the years, that have modified the functioning of the pension scheme. For the purposes of our analysis, we are primarily interested in the evolution of retirement access criteria. Consequently, we propose a summary table in Appendix Section A.2 outlining these criteria, which may not be exhaustive of minimum retirement requirements.

2.2 Pension eligibility

The key variable of the analysis is *Distance to/from eligibility* and we will refer to it as the *running variable*. This variable can be regarded as measuring relative time, taking positive values for individuals eligible for retirement and negative for those who are still waiting for eligibility. Age at eligibility is given by the minimum between the statutory age for old-age retirement and the statutory requirement in terms of years of contributions. The latter defines the so called early retirement. Thus, from an operational point of view the running variable can be computed in three steps. In the first step we compute the time to eligibility according to old-age retirement rule which is given by the difference between the individual's age at the time of the interview and his/her age at eligibility. In the second step we repeat a similar computation by using the statutory rule for early retirement, which is defined in terms of years of contributions. Finally, we take the minimum of the two.[2]

The number of years of contributions can be obtained using information from SHARE and SHARELIFE. We consider individuals as retired or working based on the employment status they declared at the time of the interview. This definition implies that the corresponding amount of declared paid work by retirees is zero.[3] Thanks to detailed information available in the survey the running variable is precisely measures in months. In principle, this allows us to avoid errors in the treatment assignment mechanism properly implementing a RDD analysis. The eligibility rules for old-age and early retirement are determined exogenously by the central government. Unfortunately, to complicate matters, a significant number of pension reforms have taken place in Italy over the last thirty years. It follows that for workers, the requirements to be taken into account are those in force at the time of the interview, whereas for retirees, reference must be made to the legislative criteria that were in force at the time of their retirement. For a general overview of the pension criteria that have succeeded in Italy over the last thirty years, see Section A.2.

For the sake of completeness, we must also mention that some exceptions to retirement rules are conceived for special jobs (particularly strenuous jobs) and/or categories of workers, e.g., the disabled. It follows that we must also consider the type of work done by individuals. For occupational categories, we use the last job category for retirees and the current occupational category for workers.

3 Data

We use data from wave 8 of the Survey of Health, Ageing and Retirement in Europe (SHARE), collected in 2019–2020. SHARE focuses on older people (over 50 years of age) and provides interesting information on different topics, such as health, financial situations, social networks, and many more. Wave 8 includes a dedicated time-use module (TE), which collects information on how respondents allocated their time during the day preceding the interview.

To reconstruct individuals' employment histories and pension eligibility conditions, we complement SHARE wave 8 with retrospective information from SHARELIFE. SHARELIFE provides detailed life-course data on education, employment spells, and retirement histories, allowing us to recover key variables such as age at first job, contribution years, and the timing of retirement. Merging the two datasets enables us to link contemporaneous time-use outcomes to institutional pension eligibility rules derived from individuals' observed work histories.

The sample consists of individuals interviewed in wave 8 who can be successfully linked to their SHARELIFE records. We restrict the sample to respondents residing in Italy and exclude

[2] Concerning the old age pension, although it is true that the main criterion is age, a minimum of contributions is always required.

[3] Notice that information contained in SHARELIFE concerns the individual's history up to 2017, whereas SHARE refers to 2019. For this reason, it was necessary to impute the years of contributions from 2017 to 2019, according to the working status of the individual. If the individual was already retired in 2017 then he/she was assigned zero years of contributions for the following two years. If the individual was found to be a worker in 2017 and in 2019 two years of contributions were assigned. Finally, if the individual is unemployed or retired in 2019, only one year of contributions has been imputed.

individuals for whom it is not possible to identify a valid retirement eligibility age. We also exclude retirees who do not report a retirement year. After applying these restrictions, the final sample includes 1,075 individuals: 636 men and 439 women. All individuals are observed either as workers or retirees at the time of the interview.

In the following subsection, we explain how we obtained the variables of interest from the dataset.

3.1 Outcome variables and descriptive statistics

The outcome variables of interest are derived from the SHARE TE module and measure the time (in minutes per day) devoted to different activities on the day preceding the interview. The data are not collected in the form of a full 24-hour diary; therefore, reported activities do not necessarily sum up to 1,440 minutes.

We focus on four main outcomes: paid work, unpaid work, total work, and leisure. Paid work measures the total time spent in market activities. Unpaid work includes time devoted to household production and caregiving activities, such as household chores, administrative tasks, and care for children, parents, partners, or other individuals. Total work is defined as the sum of paid and unpaid work. The leisure measure reflects self-reported time spent in leisure activities on the day prior to the interview.

Table 1 reports descriptive statistics for the main outcome variables (panel A) and selected predetermined covariates (panel B), separately by gender and retirement status. All time-use variables are expressed in minutes per day.

Table 1: Descriptive statistics - Male and Female

	Variable	Male		Female	
		Retired	Workers	Retired	Workers
A	Unpaid work tot	166.2	128.9	229.8	200.5
		(161.5)	(140.7)	(170.0)	(161.1)
	Paid work tot	0.0	364.1	0.0	279.8
		(0.0)	(199.5)	(0.0)	(213.5)
	Total work	166.2	493.0	229.8	480.4
		(161.5)	(193.1)	(170.0)	(216.7)
	Leisure tot	182.8	117.8	185.8	107.3
		(118.7)	(83.74)	(128.1)	(94.15)
B	Cohabiting	0.865	0.878	0.645	0.727
		(0.342)	(0.328)	(0.479)	(0.447)
	Less than primary school	0.496	0.115	0.427	0.0833
		(0.500)	(0.320)	(0.495)	(0.277)
	Primary school	0.119	0.0338	0.0977	0.0455
		(0.325)	(0.181)	(0.297)	(0.209)
	Lower high school	0.126	0.243	0.143	0.189
		(0.332)	(0.430)	(0.351)	(0.393)
	Diploma	0.230	0.480	0.264	0.492
		(0.422)	(0.501)	(0.441)	(0.502)
	University degree	0.0268	0.128	0.0530	0.189
		(0.162)	(0.336)	(0.224)	(0.393)
	In good health	0.536	0.824	0.485	0.780
		(0.499)	(0.382)	(0.501)	(0.416)
	Income household	7.478	7.711	7.474	7.770
		(0.919)	(1.119)	(0.595)	(1.152)
	Depressed	0.318	0.197	0.463	0.328
		(0.466)	(0.399)	(0.499)	(0.471)
	Household size	2.217	2.946	1.883	2.614
		(0.738)	(1.200)	(0.681)	(1.103)
	Parent alive	0.238	0.700	0.222	0.766
		(0.428)	(0.461)	(0.419)	(0.426)
	Siblings	0.899	0.903	0.871	0.944
		(0.302)	(0.296)	(0.336)	(0.230)
	Children	1.989	1.681	1.732	1.371
		(1.204)	(1.064)	(1.200)	(0.917)
	Grandchildren	2.801	1.208	2.483	0.673
		(2.621)	(1.858)	(2.288)	(1.622)
	Housewife	0.330	0.284	–	–
		(0.471)	(0.452)		

Note: mean values and standard deviations in parentesis. Panel A includes outcome variables expressed in minutes. In Panel B, all variables are dummies unless otherwise stated. Household size = number of family members.

Several patterns emerge. For both men and women, retirees spend substantially less time in total work than workers, reflecting the exit from market activities. For both genders, this reduction is driven by the sharp decline in paid work at retirement. Among women, the increase in unpaid work only partially offsets the loss of paid work, so that total work still decreases after retirement. Leisure time increases markedly after retirement for both genders, with particularly large gains among women.

Interestingly, working women appear to be more educated than working men, and for both genders workers are more educated than retirees. This pattern is consistent with cohort effects and long-run social change in educational attainment, as workers are on average younger than retired individuals.

Other notable differences emerge for health and family characteristics. Workers report significantly lower levels of depression compared to retirees and live in larger households, are more likely to have at least one living parent, and have fewer children and grandchildren. Age differences are likely a primary factor in explaining these disparities; overall, these results seem to suggest that age is one of the significant factors in explaining differences in mental health, household size, and family composition between workers and retired individuals.

This underscores the importance of adopting an identification strategy that isolates the causal effect of retirement from confounding age and cohort effects, a task we address in the following sections using a regression discontinuity design.

4 Identification strategy

4.1 The model

We estimate the effect of retirement on different components of time allocation by exploiting the discontinuity in pension eligibility introduced by the Italian laws over the years. As in the case of Battistin et al. (2009) and Eibich and Siedler (2020), the setup lends itself to a fuzzy regression discontinuity exercise.

Depending on the different reform we are considering, retirement may depend on a combination of age and years of contributions or exclusively on years of contributions. In both cases, eligibility follows a deterministic rule with fixed limits.

To fix ideas, let us denote by X_i the (normalized) running variable measuring the distance from/to the eligibility, and let $Z_i = I(X_i \geq 0)$ be the corresponding eligibility status, i.e. whenever $Z_i \geq 0$ individual i is eligible for retirement. The treatment variable, R_i, is equal to 1 for retirees and 0 otherwise. The outcomes of interest, Y_i, will be represented by some variables possibly affected by retirement status, e.g., domestic production, care work, leisure time, paid work, etc. At the cutoff (i.e. X_i=0) we will observe a discontinuous change in the probability of being treated. The probability of retirement does not jump from zero to one because retirement is not compulsory, it is an option in the hands of the worker, and some individuals may decide to continue working even if they have met their minimum pension requirements. Furthermore, eligibility rules are subject to some limited but not excludable exceptions. In particular, disabled people, war veterans, recipients of reversibility pensions, and other peculiar cases are characterized by shorter working periods compared to the common minimum requirement. In other words, it could happen that some individuals retire early, compared to the standard eligibility rule, thanks to legal exceptions. We are not able to fully control for all the factors that alter the general requirements as established by law. It follows that the setting just presented is a well-known fuzzy RDD in which the discontinuity in the probability of retirement helps to overcome the problems arising from endogenous selection in retirement. For a recent and detailed guidance to practical implementation of RDD we refer the reader to Cattaneo et al. (2024a) and Cattaneo et al. (2024b). Specifically, our setting is a fuzzy case with double non-compliance, that is, on the one side we have ineligible retirees (e.g. disabled people, veterans, etc.) who are actually retired, and on the other side we have eligible individuals who are still working. In this setup RD can recover the local average treatment effect (LATE), i.e. the effect on compilers at the cutoff, where compliers are defined as individuals whose treatment status changes as we move the value of X_i from just to the left of the cutoff to just to the right of it (Angrist and Pischke, 2009). This is like saying that compliers are individuals who receive treatment if their score is above the threshold but do not receive it if their score is below the threshold.

A valid fuzzy RD design relies on three main assumptions. The first assumption requires that conditional on R there is *no discontinuity* in the regression functions at the cutoff. This implies that in the absence of retirement, the allocation of time should not change at the cutoff point. In other words, "all other factors" driving time allocation must be continuous at the threshold (see, e.g. Hahn et al., 2001).

The second assumption requires that the potential treatment $R_i(X)$ must be *monotonic* with respect to the score X at the threshold c. This implies that there are no "defiers," - that is , individuals who would have received the treatment if their score were below the threshold but would not have received it if their score were above the threshold; and there is at least one complier.

The third assumption, not always explicitly stated, is that there is a discontinuity in the probability of receiving treatment at the cutoff, commonly referred to as the existence of a *first stage*. This means that the treatment probability must show a jump at the cutoff.

From a strictly empirical point of view, RDD is implemented using instrumental variable (IV) regression equations, where the instrument is eligibility Z_i. The first stage equation is as follows:

$$R_i = \alpha_1 + \theta Z_i + \beta_1 f\left(X_i\right) + \epsilon_{1i} \tag{1}$$

where $f(X_i)$ incorporates a power series of X_i and its interaction with Z_i to allow for different slopes on either side of the cutoff, and θ captures the jump in the probability of retirement induced by achieving eligibility, i.e. the proportion of compliers in the sample, ϵ_{1i} is a random error term.

The reduced form equation is:

$$Y_i = \alpha_2 + \lambda Z_i + \beta_2 f\left(X_i\right) + \epsilon_{2i} \tag{2}$$

where λ is the causal effect of eligibility, Z on Y and ϵ_{2i} is an random error term. Other controls can be easily added to equations, (3.1) and (3.2), and the causal effect of R on Y for compliers is obtained as the ratio $\frac{\lambda}{\theta}$.

The IV exercise is carried out at the cutoff, meaning in a narrow bandwidth around the cutoff. The amplitude of the bandwidth is an empirical question, and we will test the robustness of our choice. Table 2 reports further descriptive statistics for the key variables used in our RDD design, separately for men and women. The average age at interview is slightly higher for men than for women (approximately 72.6 vs. 70.3 years). The standard deviation is slightly higher for women, suggesting slightly more variation among women. This age difference is reflected in the distribution of the running variable. On average, men are further away from the eligibility threshold than women, although the dispersion of the running variable is very similar across genders.

Consistently with these patterns, a higher share of men are eligible for retirement at the time of the interview (82.5 percent versus 75.2 percent among women) and a higher fraction are already retired (76.7 percent versus 69.9 percent). These differences reflect gender-specific career trajectories and institutional rules governing retirement eligibility in Italy, which generate distinct distributions of eligibility and retirement status by gender.

Table 2: Further descriptive statistics

	Males	Females
Age	870.8	843.5
	(105.5)	(113.5)
time to/from eligibility (X)	190.4	134.6
	(169.1)	(171.6)
Eligibility variable (Z)	0.825	0.752
	(0.380)	(0.433)
Treatment variable (R)	0.767	0.699
	(0.423)	(0.459)

Note: mean values and standard deviations in parenthesis. Age and time to/from eligibility are measured in months.

4.2 Testing the assumptions

The first assumption, continuity of the conditional regression of the expected outcome, can be indirectly tested by checking whether the relevant variables change significantly at the cutoff. The reasoning is that if covariates that are known to be strongly correlated with the outcome of interest are discontinuous at the cutoff, the continuity of the potential outcome functions is unlike to hold (Calonico et al., 2019). Put differently, if we find causal effects on pre-determined covariates that clearly must not be affected by retirement, then the identification breaks down. Briefly, the test is conducted by using as outcome variables each covariate listed in the columns of Table 3. According to Cattaneo et al. (2024a) for falsification purposes, the optimal Coverage Error Rate (CER) bandwidth is the appropriate data-driven criterion, since we are interested in inference and not in the point estimates.

Table 3: Falsification test by gender

	Males		Females	
Variables	Effect	Robust P-val	Effect	Robust P-val
Cohabiting	0.020	0.856	-0.864	0.103
Less than primary school	0.093	0.903	-0.180	0.631
Primary school	0.028	0.590	0.093	0.509
Lower high school	-0.047	0.967	-0.002	0.968
Diploma	-0.042	0.858	-0.106	0.906
University degree	0.107	0.261	0.046	0.950
In good health	-0.138	0.839	-0.291	0.449
Income household	-0.713	0.305	0.391	0.649
Depressed	-0.055	0.880	0.427	0.473
Household size	0.320	0.766	-1.333	0.173
Parent alive	-1.148	0.428	0.599	0.232
Siblings	-0.235	0.381	0.104	0.726
Children	-1.047	0.399	-1.485	0.134
Grandchildren	-0.572	0.809	0.061	0.873
Housewife	0.131	0.713	0.336	0.297

Note: The table reports the falsification test according to the CER optimal BW selector. Heteroskedasticity Robust P-values are used.

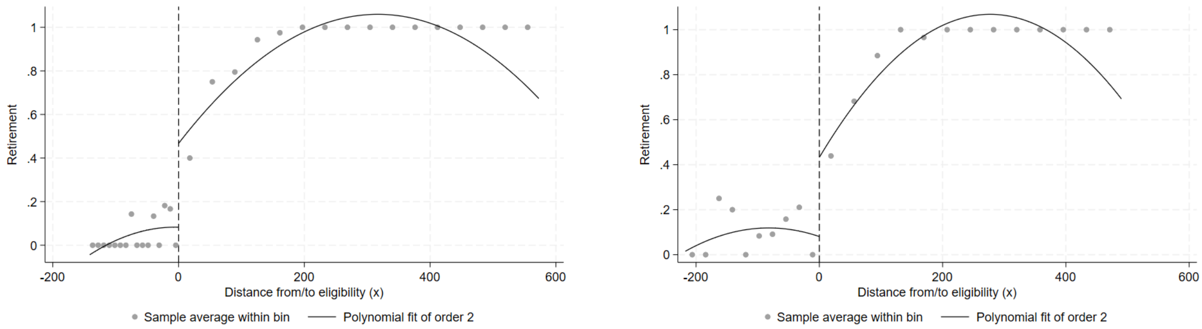
The results show no statistically significant discontinuities at conventional levels for all the covariates analyzed, both for men and women. This finding supports the validity of the design and suggests that individuals just below and just above the eligibility threshold are comparable in terms of observable characteristics.

An issue strictly related to *continuity* is manipulation of the running variable. Indeed, if manipulation takes place, there will no longer be randomly assigned treated/controls at the cutoff and the running variable will not be distributed continuously. The manipulation hypothesis is actually unlikely in our context because to qualify for retirement individuals should manipulate either their age or the amount of contributions paid, or both. It is evident that age cannot be changed, and it is also hard to believe that contribution periods can be manipulated, given that these are recorded and calculated by the National Social Security Institute, INPS. Notwithstanding, a formal test of manipulation can be carried out by following the approach by McCrary (2008), that, in line with our expectation, returns a P-Value of the null hypothesis of no manipulation equal to 0.3763 for men and 0.2906 for women. Visual evidence for the distribution of the running variable along with the McCrary (2008) test are given in Figures A.1

and A.2.

As far as the *first stage* assumption is concerned, Figure 1 graphically proves the existence of a first stage, i.e. the fact that there exists a positive jump in the probability of being treated (retired) for eligible individuals.

Figure 1: First stage plot by gender



(**a**) First Stage - Male

(**b**) First Stage - Female

Note: The figure reports the jump in the probability of treatment at the cutoff. On the y-axis of each panel there is the probability of retirement, while the x-axis reports time from/to eligibility, centered at zero. The black line represents the underlying regression function, the dots represent the average probability at a given relative time, the dashed line is the cutoff. The bins are computed according to the evenly-spaced method as proposed by Calonico et al. (2015).

5 Results

5.1 First Stages

We move on now to the estimation of the parameters of the causal effect of Z_i on R_i; namely, the first stage that provides us with relevant information concerning the effect of eligibility on retirement at the cutoff, as well as the proportion of compliers and the take-up rate of the policy. Following Porter (2003) and Gelman and Imbens (2019) our preferred benchmark is the quadratic specification.[4] Table 4 reports the benchmark, while, for the sake of robustness, a variety of alternative specifications are reported in Table A.1 for men and Table A.2 for women.

[4] Porter (2003) claims that local polynomials of lower order than quadratics suffer from boundary bias, while Gelman and Imbens (2019) show that local polynomials of higher order than quadratics tend to be inaccurate.

Table 4: First stage by gender

		avg. ret. pre-thre.	Males	avg. ret. pre-thre.	Females
A	Eligibility dummy (Z)		0.588*** (0.169)		0.680*** (0.169)
	Obs	0.075	143	0.137	104
B	Eligibility dummy (Z)		0.608*** (0.172)		0.671*** (0.169)
	Obs	0.081	131	0.135	109
C	Eligibility dummy (Z)		0.604*** (0.172)		0.677*** (0.169)
	Obs	0.080	132	0.137	106
D	Eligibility dummy (Z)		0.595*** (0.169)		0.667*** (0.168)
	Obs	0.076	141	0.130	113

Note: Heteroskedasticity robust standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1. The table reports the probability discontinuity (i.e., the regression of R on Z). Row A identifies the first stages of the outcome variable 'Unpaid work'; the subsequent rows refer to: 'Paid work', 'Leisure', and 'Total work', respectively. The columns 'avg. ret.' 'pre-thre.' report the pre-threshold average retirement rate computed within the bandwidth.

According to the notation of eq.(3.1) the tables are expected to report unique estimates of θ for each specification, which implies that Table 4 should contain only one row, regardless of the outcome, Y. However, the outcome variables exhibit different missing value patterns, necessitating a repetition of the first-stage estimation to account for these missing values.

The estimates are positive and statistically significant across all the specifications and for both men and women. Comparing more closely the results by gender, the magnitude of the coefficients is large for both groups, ranging between about 60% and 70%, with values that are consistently slightly higher for women across the different specifications. The pre-threshold average retirement rate (computed within the optimal bandwidth) is around 7–8% for males and approximately 13–14% for females. This is a first interesting result showing that retirement take-up at eligibility is substantial for both genders, while women exhibit somewhat higher baseline retirement rates before the threshold.

In the Italian context, this pattern is likely to reflect institutional and life-course mechanisms that differentially affect women's retirement behavior. These include the availability of gender-specific early retirement schemes (such as *Opzione Donna*) and career patterns characterized by more discontinuous employment histories and caregiving responsibilities, which may induce women to exit the labor market earlier, even before full statutory eligibility is reached. More broadly, these findings are consistent with evidence showing that women tend to be more responsive than men to changes in pension incentives, partly because their labour supply is typically more elastic and less strongly attached to the labour market (see, e.g., Carta and De Philippis (2024)). As a result, women may display higher baseline retirement rates below the cutoff, while still responding strongly to eligibility once the threshold is crossed.

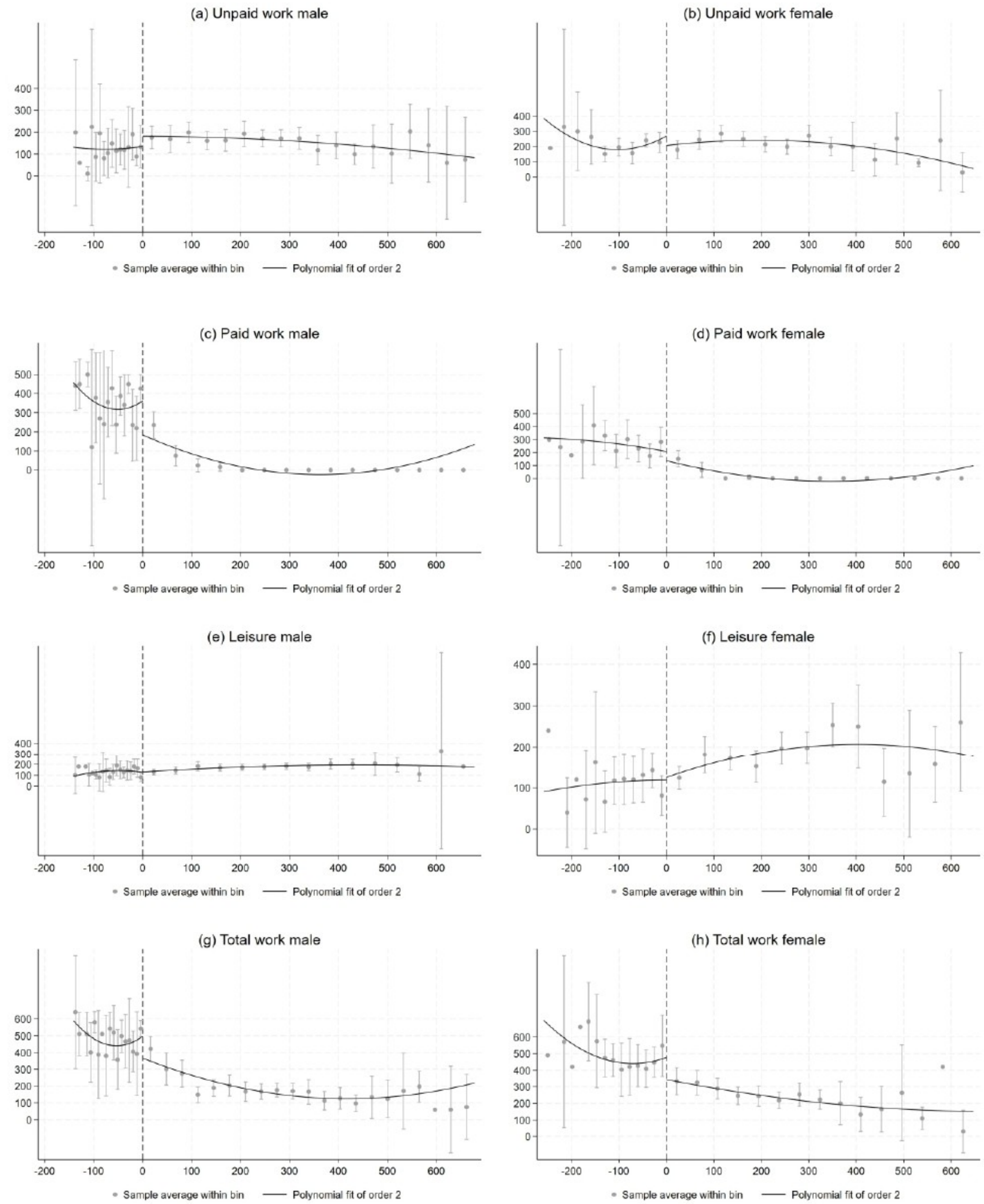
Our estimates for men are higher than the one of Battistin et al. (2009) for Italy, who found a value of around 43% on a sample covering the years 1993-2004. The increase over time is in line with pension reforms that have consistently eroded workers' benefits. In addition, we can safely claim that it is a consequence of the ongoing trend of raising the retirement age and lengthening contribution periods (see Section A.2).

This parameter may have clear relevant implications in terms of policy, and we can also look at this result from another vantage point. Indeed, the first stage represents the share of compliers in the sample, which, in turn, is nothing but the take-up rate of the policy; this gives rise to two considerations. First, a high first stage implies that pension reforms play a relevant role, affecting a large share of the population. Second, if a small perturbation of the cutoff leads to a substantial change in the first stage, the complier population changes dramatically with small changes in the eligibility. As a consequence, policy reforms that slightly alter the eligibility cutoff may have a great impact on retirees. Given the policy relevance of this issue, we will delve into it further in Section 6 where formal tests will demonstrate the stability of the first stage parameter for both men and women.

5.2 Fuzzy RD parameters

Again, we first present a visual analysis of the effect under scrutiny, and then we move on to a more formal approach. Figure 2 reports the jump in outcomes due to eligibility. The results for each outcome variable are reported for men on the left and for women on the right. Each point in the figures represents the average time (measured in minutes per day) spent on a given activity in a given relative time; the bold vertical line is the cutoff, while the vertical gray lines represent the 95% confidence intervals.

Figure 2: RD plot: jump in the outcomes at the cutoff



Note: the figure reports the jump in the outcomes due to eligibility. The left-hand-side reports the results for each outcome variable for men, while the corresponding graph for women are on the on the right-hand-side. On the y-axis of each panel time is measured in minutes, while the x-axis reports time from/to eligibility (measured in months), centered at zero. The black line represents the underlying regression function, the dots represent the average time spent on a given activity in a given relative time, the vertical bold line is the cutoff, while the vertical gray lines represent the 95% confidence intervals. The bins are computed according to the evenly-spaced method as proposed by Calonico et al. (2015).

Visual inspection of the jump provides a useful anticipation of the estimates and some common features can be detected from the plot in Figure 2. Interestingly, a positive jump in unpaid work is visible for men, while a negative jump emerges for women. This gender asymmetry already provides interesting descriptive evidence on how retirement affects the allocation of unpaid activities.

When it comes to paid work, a clear negative jump is visible for both men and women, although the drop is less pronounced for women. This pattern is reflected in a reduction in total work at the eligibility threshold for both genders. With regard to leisure, no clear increase appears for men, while women show an increase around the eligibility threshold.

Data fit considerations are noteworthy. Specifically, leisure hours tend to increase as individuals move further past the eligibility threshold. This pattern suggests a progressive engagement in new activities during retirement, a finding consistent with the fitted leisure profiles for both men and women.

Table 5 reports the RD estimates by gender, with and without covariate adjustment. In addition, Tables A.3 - A.6 in the Appendix report a series of robustness checks.

The econometric results for men in Table 5 show statistically significant effects of retirement on paid work. In all specifications, retirement leads to a large and significant reduction in paid work that is robust to the inclusion of covariates. Regarding the magnitude of the coefficients, the estimated decline corresponds to a substantial reduction in daily paid work (roughly $5-6$ hours). This translates into a significant reduction in total work, which is marginally significant in the baseline specification and becomes strongly significant once covariates are accounted for. Simultaneously, retirement is associated with a statistically significant increase in leisure for men, an effect that holds even when controlling for covariates. The estimated increase of $2-3$ hours represents a substantial rise in daily leisure time. No statistically significant effects are detected for unpaid work. These patterns are confirmed across the alternative specifications in Tables A.5 and A.6. Taken together, these results indicate a reallocation of time away from market work toward leisure among men at retirement, whereas for women the adjustment is more muted and less precisely estimated across time-use margins.

For women, retirement is associated with a reduction in total work once covariates are included, roughly 6 hours a day. This pattern is consistent with a decline in market-related activities, even though the estimates for paid work alone are not statistically significant across specifications. At the same time, the point estimates for unpaid work are negative, suggesting a partial offset of market work reductions through adjustments in non-market activities, although these effects are not precisely estimated. Finally, retirement is associated with a contained increase in leisure, $1-2$ hours. These patterns are stable across alternative specifications reported in Tables A.3 and A.4.

The findings of our study align well with conventional economic models that consider individuals facing a trade-off between paid work and other activities due to the opportunity costs associated with time investment. With a finite amount of time available, time must be balanced between market-based activities and non-market activities. This is the so-called 'activity substitution' hypothesis.

Before retirement, most people prioritise paid work to meet consumption needs, making the opportunity cost of engaging in other pursuits relatively high. However, retirement alters this dynamic: with consumption needs covered by pension income, individuals experience reduced time constraints. As a result, retirees have more freedom to allocate their time to leisure, unpaid work, and other activities.

Previous research has already found that retirement affects how people choose to spend their time (Bergeot and Fontaine, 2020; Eibich and Siedler, 2020; Bonsang and Van Soest, 2020; Ciani, 2016) but these studies found the highest effect for unpaid work, with a limited impact for men and a strong positive effect for women. To the best of our knowledge, evidence on leisure as the main margin of adjustment remains scant, or even absent, in the literature.

In contrast, our results suggest that the main adjustment following retirement operates through a reduction in paid work and total work, especially for men, with leisure emerging as a key margin of time reallocation. For men, the increase in leisure corresponds to additional minutes per day devoted to leisure activities, indicating that a substantial share of the time released from paid work is absorbed by leisure. For women, a reduction in total work is also observed, although the adjustment across specific time-use categories is more gradual and less precisely estimated. Overall, the evidence indicates that leisure plays an important role in absorbing the time freed up by retirement.

While gender differences in the magnitude and precision of the estimates remain, these patterns are consistent with a potential narrowing of the gender gap in time allocation at older ages. This evolution may reflect broader changes in social norms, labour market attachment, and life-course trajectories, which could have contributed to reshaping how men and women allocate their time after exiting the labour market.

Table 5: RD estimates by gender and covariates adjustment

Variables	Females w/o covs	Females with covs	Males w/o covs	Males with covs
Unpaid Work	-144.3	-187.0	159.5	-147.0
Robust Bias-Corr. SE	(122.90)	(136.05)	(125.74)	(106.16)
Robust Bias-Corr. p-value	[0.267]	[0.160]	[0.195]	[0.194]
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	mserd	mserd	mserd	mserd
Order Local Polyn	2	2	2	2
BW	±62.714	±63.418	±63.683	±73.126
Paid Work	-103.6	-150.3	-385.0**	-296.6**
Robust Bias-Corr. SE	(167.36)	(168.81)	(84.87)	(122.61)
Robust Bias-Corr. p-value	[0.532]	[0.442]	[0.000]	[0.020]
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	mserd	mserd	mserd	mserd
Order Local Polyn	2	2	2	2
BW	±64.513	±68.696	±57.984	±61.766
Leisure	115.2*	61.56	119.8*	203.1*
Robust Bias-Corr. SE	(70.05)	(69.89)	(56.74)	(99.93)
Robust Bias-Corr. p-value	[0.072]	[0.455]	[0.031]	[0.055]
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	mserd	mserd	mserd	mserd
Order Local Polyn	2	2	2	2
BW	±63.377	±60.608	±59.390	±76.314
Total Work	-252.8	-351.0**	-203.0	-379.4***
Robust Bias-Corr. SE	(171.78)	(189.10)	(131.01)	(90.93)
Robust Bias-Corr. p-value	[0.154]	[0.081]	[0.095]	[0.001]
Kernel Type	Triangular	Triangular	Triangular	Triangular
BW Type	mserd	mserd	mserd	mserd
Order Local Polyn	2	2	2	2
BW	±65.483	±75.312	±62.259	±59.951

Note: The outcome variables (measured in minutes) are reported in the first column in bold. Bias-corrected robust std errors in parentheses. ***p<0.01; **p<0.05; *p<0.10.

6 External validity

One of the drawbacks of RDD is that the results apply only at the cutoff and cannot be extended to other populations. In this section, we follow the approach proposed by Dong and Lewbel (2015) to assess the external validity of our results. The authors show that the derivative of the treatment effect with respect to the running variable at the cutoff (TED) has significant implications for extrapolating the estimated LATE beyond the cutoff. Specifically, a large TED raises concerns about the external validity of the estimated LATE. Conversely, if the identified TED is locally zero, the estimated LATE is more likely to exhibit higher external validity. Table 6 reports the estimated TED's by gender.

All TED estimates are statistically insignificant at conventional levels, with the exception of a marginally significant effect for leisure among women. Overall, these results imply stability of the estimates in the sense that small changes in the cutoff do not substantially alter the estimated RD effect, thereby increasing confidence in the validity of the LATE beyond the exact cutoff.

Closely connected to TED is the complier probability derivative (CPD) which measures the stability of the compliers population in fuzzy designs. Having a CPD near zero suggests the potential stability of the compliers population, whereas a large value (positive or negative) of the CPD means that the compliers population changes dramatically with small changes in X.

Table 6: TED and CPD Estimates by Gender

Outcome variable	TED Male		TED Female		CPD Male		CPD Female	
	TED	P-val	TED	P-val	CPD	P-val	CPD	P-val
Unpaid work	2.14	0.790	6.49	0.800	0.005	0.210	0.009	0.104
Paid work	-9.23	0.354	-7.79	0.263	0.005	0.295	0.009*	0.087
Leisure	0.09	0.986	9.83*	0.070	0.005	0.277	0.009*	0.099
Total work	-7.47	0.491	-2.13	0.796	0.006	0.207	0.009*	0.087

Note: The table reports the estimated TED and CPD and their heteroskedasticity robust P-values for each outcome. ***p<0.01, **p<0.05, *p<0.1.

The last four columns of Table 6 report the CPD estimates, which are generally small and mostly statistically insignificant across outcomes and genders. Marginal statistical significance emerges for women in paid work, leisure, and total work, with CPD estimates close to 0.01 and p-values near conventional thresholds. Importantly, the associated magnitudes are very limited, amounting to roughly 1% of the corresponding first-stage effect, suggesting that any changes in the complier population around the cutoff are quantitatively small.

Taken together, the TED and CPD results provide no evidence of substantial sensitivity of either the treatment effect or the share of compliers to small shifts in eligibility. This pattern supports the interpretation that the estimated fuzzy RD effects are not driven by local instability in compliance behavior near the cutoff.

7 Conclusions and policy implications

This paper examines gender differences in time reallocation following retirement in Italy, providing causal evidence on how men and women adjust paid work, unpaid work, leisure, and total work at the retirement threshold.

A large body of literature documents substantial gender inequalities in labor market participation and time allocation during working life, with women bearing a disproportionate share of unpaid work and caregiving responsibilities and enjoying less leisure time than men. However, much less is known about how these disparities evolve after retirement. Existing studies typically find an increase in unpaid work among retired women, with limited evidence of gains in leisure time.

This paper contributes to the literature by exploiting institutional discontinuities in pension eligibility to identify the causal effect of retirement on time use. Using a fuzzy RDD and detailed Italian time-use data, we show that retirement induces a significant reduction in total work for both men and women, although the underlying margins of adjustment differ by gender.

For men, the reduction in total work is primarily driven by a sharp decline in paid work, accompanied by a robust increase in leisure time. For women, total work also declines at retirement, but the adjustment is more gradual and spread across different time-use categories, with a reduction in unpaid work and an increase in leisure.

These findings highlight that retirement does not simply reinforce pre-existing gender roles in time allocation. Instead, they suggest that women experience a rebalancing of time away from unpaid work and toward leisure, contrasting with earlier evidence for Italy and other countries. In this respect, our results point to important changes in post-retirement behavior across cohorts.

An additional implication of our findings concerns the role of leisure time after retirement. From a public policy perspective, this pattern is highly relevant, as leisure represents a key dimension through which retirement may affect individual well-being and ageing trajectories. In the health and ageing literature, leisure activities such as physical exercise, social interaction, and cognitive engagement are widely recognised as important determinants of healthy ageing, defined as the ability to maintain functional capacity and well-being in later life (World Health Organization, 2021; Zhang and Zhang, 2019).

While this paper does not directly assess health outcomes, the observed increase in leisure highlights the importance of understanding how retirees use their newly available time. A growing literature in health economics suggests that different forms of leisure - ranging from physically and socially active pursuits to more sedentary activities - can have markedly different implications for physical health, cognitive functioning, and healthcare utilization at older ages. In this respect, our results point to time use as a potentially important channel linking retirement behavior to long-run public health expenditures and fiscal sustainability, and call for further research on the quality and composition of leisure after retirement. Besides, our findings indicate that statutory retirement eligibility rules play a central role in shaping retirement behavior and time allocation for both men and women. High take-up rates at the eligibility threshold imply that even relatively small changes in eligibility conditions can have sizable effects on retirement decisions and on how individuals allocate their time after leaving the labor market.

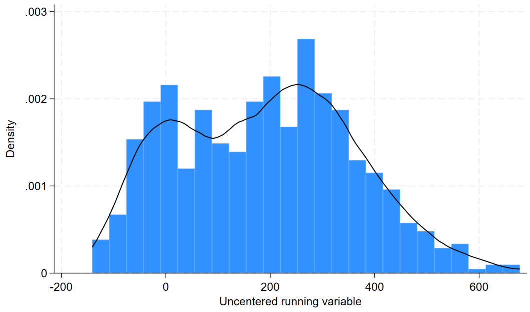
While gender differences in the magnitude and precision of the estimated effects remain, the observed patterns are consistent with a potential narrowing of gender gaps in time allocation at older ages. This evolution may reflect broader changes in labor market attachment, family roles, and social norms across cohorts.

Finally, this study is subject to some limitations. Leisure time aggregates activities with heterogeneous implications, and gender-specific reporting differences may affect measurement. Moreover, while we distinguish individuals by cohabitation status, we do not explore more detailed marital trajectories or occupational heterogeneity, which remain important avenues for future research using larger or administrative data. We leave these issues to further research should appropriate data become available.

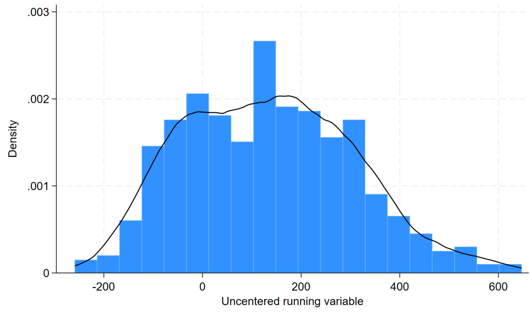
Appendix

A Further evidence

Figure A.1: RD Plot: Kernel Density

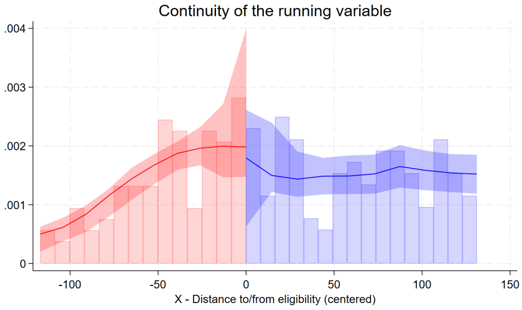


(**a**) Males

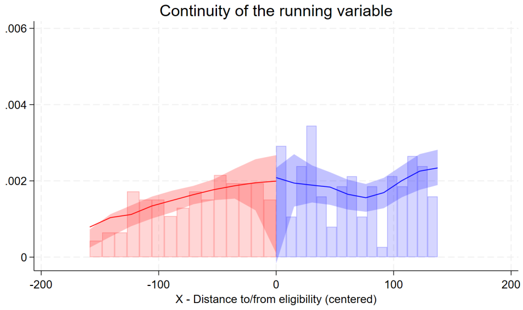


(**b**) Females

Figure A.2: Manipulation test



(**a**) Males



(**b**) Females

A.1 Tables

Table A.1: First Stage - Males

	Variables	(1) Linear	(2) Asymmetric BW	(3) Uniform Kernel	(4) Cluster SE
A	Eligibility dummy (Z)	0.448***	0.366***	0.672***	0.588***
		(0.155)	(0.093)	(0.185)	(0.165)
	Obs	118	423	101	143
B	Eligibility dummy (Z)	0.500***	0.371***	0.627***	0.612***
		(0.159)	(0.099)	(0.182)	(0.166)
	Obs	101	399	104	131
C	Eligibility dummy (Z)	0.488***	0.365***	0.652***	0.607***
		(0.159)	(0.100)	(0.189)	(0.166)
	Obs	104	370	98	132
D	Eligibility dummy (Z)	0.461***	0.368***	0.591***	0.594***
		(0.155)	(0.099)	(0.171)	(0.166)
	Obs	113	389	119	141

Note: Heteroskedasticity robust standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1. The table reports the probability discontinuity (i.e., the regression of R on Z). Row A identifies the first stage of the outcome variable 'Unpaid work'; rows B–D refer to: 'Paid work', 'Leisure', and 'Total work', respectively.

Table A.2: First Stage - Females

	Variables	(1) Linear	(2) Asymmetric BW	(3) Uniform Kernel	(4) Cluster SE
	Eligibility	0.561***	0.418***	0.68***	0.665***
A	dummy (Z)	(0.144)	(0.128)	(0.169)	(0.168)
	Obs	93	256	104	114
	Eligibility	0.502***	0.417***	0.671***	0.663***
B	dummy (Z)	(0.135)	(0.12)	(0.169)	(0.168)
	Obs	113	305	109	116
	Eligibility	0.552***	0.5***	0.677***	0.664***
C	dummy (Z)	(0.145)	(0.143)	(0.169)	(0.168)
	Obs	94	185	106	114
	Eligibility	0.567***	0.412**	0.667***	0.661***
D	dummy (Z)	(0.145)	(0.125)	(0.168)	(0.168)
	Obs	93	287	113	118

Note: Heteroskedasticity robust standard errors in parentheses, ***p<0.01, **p<0.05, *p<0.1. The table reports the probability discontinuity (i.e., the regression of R on Z). Row A identifies the first stage of the outcome variable 'Unpaid work'; rows B–D refer to: 'Paid work', 'Leisure', and 'Total work', respectively.

Table A.3: RD Estimates for females

Variables	(1) Linear	(2) Asymmetric BW	(3) Triangular Kernel	(4) Cluster SE
RD estimates for **Unpaid Work**	-157.8	-239.1	-93.77	-146.5
Robust Bias-Corr. SE	(125.62)	(215.99)	(116.77)	(114.24)
Robust Bias-Corr. p-value	[0.253]	[0.266]	[0.476]	[0.230]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±54.830	294.308	±73.730	±66.091
RD estimates for **Paid Work**	-159.9	-145.8	-175.9	-109.9
Robust Bias-Corr. SE	(165.41)	(237.78)	(152.15)	(175.71)
Robust Bias-Corr. p-value	[0.310]	[0.592]	[0.215]	[0.529]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±65.475	366.207	±57.429	±67.037
RD estimates for **Leisure**	87.12	129.8	159.5**	118.3
Robust Bias-Corr. SE	(66.41)	(96.05)	(82.03)	(85.80)
Robust Bias-Corr. p-value	[0.147]	[0.183]	[0.036]	[0.126]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±56.704	211.971	±61.266	±66.508
RD estimates for **Total Work**	-313.7**	-379.3	-265.2*	-257.7
Robust Bias-Corr. SE	(170.23)	(270.63)	(168.11)	(181.54)
Robust Bias-Corr. p-value	[0.076]	[0.152]	[0.120]	[0.171]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±53.787	337.759	±55.296	±68.148

The dependent variables are those highlighted in bold.

*Heteroskedasticity robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.10*

Table A.4: RD Estimates for females - with covariates

Variables	(1) Linear	(2) Asymmetric BW	(3) Triangular Kernel	(4) Cluster SE
RD estimates for **Unpaid Work**	-188.6	-155.6	-175.0	-182.4*
Robust Bias-Corr. SE	(134.04)	(194.33)	(129.82)	(118.22)
Robust Bias-Corr. p-value	[0.145]	[0.424]	[0.215]	[0.117]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±58.627	62.950	±56.941	±61.482
RD estimates for **Paid Work**	-183.3	-193.8	-203.0	-144.4
Robust Bias-Corr. SE	(159.61)	(198.41)	(182.05)	(142.34)
Robust Bias-Corr. p-value	[0.269]	[0.357]	[0.346]	[0.371]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±64.149	67.834	±61.221	±65.731
RD estimates for **Leisure**	50.89	27.52	136.0**	70.07
Robust Bias-Corr. SE	(67.62)	(80.43)	(67.74)	(78.80)
Robust Bias-Corr. p-value	[0.257]	[0.752]	[0.024]	[0.426]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±58.286	59.385	±59.570	±62.077
RD estimates for **Total Work**	-375.5**	-371.9*	-403.4**	-349.7**
Robust Bias-Corr. SE	(193.17)	(243.31)	(212.61)	(171.43)
Robust Bias-Corr. p-value	[0.053]	[0.148]	[0.049]	[0.055]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±70.043	74.785	±65.184	±72.913

The dependent variables are those highlighted in bold.

*Heteroskedasticity robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.10*

Table A.5: RD Estimates for males

Variables	(1) Linear	(2) Asymmetric BW	(3) Uniform Kernel	(4) Cluster SE
RD estimates for **Unpaid Work**	126.3	243.3*	201.4*	159.3*
Robust Bias-Corr. SE	(144.70)	(155.64)	(124.99)	(99.87)
Robust Bias-Corr. p-value	[0.389]	[0.078]	[0.068]	[0.098]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±50.877	62.608	±41.534	±63.729
RD estimates for **Paid Work**	-340.6**	-505.6**	-390.7**	-387.5**
Robust Bias-Corr. SE	(103.93)	(134.94)	(90.97)	(98.98)
Robust Bias-Corr. p-value	[0.000]	[0.000]	[0.000]	[0.000]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±41.326	56.714	±42.971	±57.205
RD estimates for **Leisure**	106.5	121.5	130.0*	120.2*
Robust Bias-Corr. SE	(62.52)	(81.76)	(58.67)	(55.87)
Robust Bias-Corr. p-value	[0.053]	[0.093]	[0.033]	[0.029]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±42.897	57.697	±39.830	±58.420
RD estimates for **Total Work**	-179.8	-236.9**	-157.1	-203.5**
Robust Bias-Corr. SE	(158.13)	(136.63)	(144.24)	(99.98)
Robust Bias-Corr. p-value	[0.181]	[0.079]	[0.238]	[0.028]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±48.429	57.470	±51.329	±62.474

The dependent variables are those highlighted in bold.

*Biased-corrected robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.10*

Table A.6: RD Estimates for males - with covariates

Variables	(1) Linear	(2) Asymmetric BW	(3) Uniform Kernel	(4) Cluster SE
RD estimates for **Unpaid Work**	-125.5	3.273	-173.9	-154.1
Robust Bias-Corr. SE	(165.01)	(151.24)	(139.41)	(133.72)
Robust Bias-Corr. p-value	[0.316]	[0.950]	[0.219]	[0.293]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±66.006	74.117	±61.972	±75.880
RD estimates for **Paid Work**	-339.9**	-409.8**	-280.9**	-296.8**
Robust Bias-Corr. SE	(164.59)	(156.98)	(137.07)	(129.80)
Robust Bias-Corr. p-value	[0.043]	[0.011]	[0.086]	[0.028]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±57.017	62.003	±50.180	±61.973
RD estimates for **Leisure**	209.1**	107.3	193.6**	199.1**
Robust Bias-Corr. SE	(124.96)	(84.59)	(82.91)	(75.25)
Robust Bias-Corr. p-value	[0.067]	[0.176]	[0.030]	[0.012]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±57.582	67.609	±44.940	±74.733
RD estimates for **Total Work**	-486.0***	-314.6**	-354.0***	-396.9**
Robust Bias-Corr. SE	(176.81)	(131.07)	(116.14)	(89.51)
Robust Bias-Corr. p-value	[0.003]	[0.012]	[0.030]	[0.000]
Kernel Type	Triangular	Triangular	Uniform	Triangular
BW Type	mserd	msetwo	mserd	mserd
Order Local Polyn	1	2	2	2
BW	±71.021	56.771	±50.835	±62.903

The dependent variables are those highlighted in bold.

*Biased-corrected robust standard errors in parentheses. ***p<0.01; **p<0.05; *p<0.10*

A.2 Pension Requirements

As regards workers, as explained in the text, the retirement criteria are those in force in the years 2019 and 2020. These years in fact coincide with the period in which the questionnaire was submitted to the subjects who, at the time of the interview, declared themselves as workers. In particular, the criteria in force for workers in the 2019-2020 period could concern the old-age pension or early-retirement pension (formerly seniority pension). According to the criteria of the old-age pension, the worker could retire at the age of 67 (also respecting a minimum number of contribution years). To be able to access the early-retirement pension, alternatively, the individual had to comply with an uniquely contributory criterion which required having at least 42 years and 10 months of contributions[5].

As regards pensioners, however, the criteria to be taken into account concern the laws in force in the years of retirement declared by the pensioners when completing the questionnaire.

Although there have been numerous pension reforms that have occurred over the last fifty years, the major changes have been introduced since the 1990s[6].

The reforms implemented from 1969 until the late 1980s occurred at a time when the Italian economy was growing, and there was no concern about the old-age crisis and improvements in longevity that undermined the system starting from the late 1980s.

The main reforms of the social security system took place in 1992, 1995, 2007 and 2011:

- the "Amato Reform" of 1992 (Legislative Decree n.503/1992) explicitly distinguishes between workers with at least 15 years of contributions by the end of 1992 and all other workers. The old system (introduced in 1969) applied, with some modifications, to the former, while the new system applies only to the latter. Additionally, the age requirement for an old-age pension was to be gradually increased by one year every two years, starting from 1994, until reaching 65 years for men and 60 years for women by 2000. Starting from 1994, for both men and women, the requirement is set at the age of 65 for state employees and 60 years for employees of local administrations. The old requirements remain unchanged for some special categories (military and police personnel, flight personnel, traveling personnel of public transportation services, firefighters, and employees of the entertainment industry). The number of years of contributions required for an old-age pension is gradually increased by one every two years starting from 1993, until reaching 20 years of contributions in 2001.

- The "Dini Reform" of 1995 (Law n.335/1995) changed the pension system from a defined benefit (DB) calculation method to a notionally defined contribution (NDC) method. The reform provided for a transitional phase during the which benefits would be calculated in proportion to the number of years of contributions under the two regimes (contributions

[5]There was also a third possibility of retirement (always falling within the category of early retirement) which concerned individuals who were 100% covered by the contributory system (and therefore had no contributions prior to January '96). Although not explicitly described in the text, this additional criterion was taken into account during the construction of the running variable.

[6]For a more extensive and techincal description of the historical and institutional aspects of the Italian social security system refer to (Brugiavini et al., 2021).

made after 1993 are covered by the new regime). During the transitional phase, a common requirement of age and years of contributions was introduced for workers to be eligible for old-age pension: from 1996, a worker could retire if they were 52 years old and had accumulated 35 years of contributions. These limits were gradually increased so that by 2002, the entry age was 57 years with 35 years of contribution, for both men and women. For public sector workers, both men and women, the rules are similar.The minimum age requirements for pensions were gradually increased to reach 40 years by 2008. Under the new regime (depending on the number of years of contributions collected), workers, both men and women, who started working after 1995 can choose the retirement age between 57 and 65 years, with a minimum requirement of at least 5 years of contributions.

- the "Damiano-Padoa-Schioppa Reform" (Law n.247/2007) raises the age criterion for workers who started working after 1995. Women are entitled to an old-age pension when they reach 65, men when they reach 60, both must have at least 5 years of paid contributions. At any age it is possible to retire with at least 40 years of contributions. From this moment on, even those who started working after 1995 can access the old-age pension. As regards early retirement: in 2008 and 2009 the joint requirement for retirement was the age of 58 with at least 35 years of contributions; regardless of age with at least 40 years of contributions.
  Since July 2009, eligibility depends on an index which is the sum of the worker's age and the years of contributions, the so-called "quota system" (in 2009 the quota was 95: 59 years of age with 36 of contributions. From 2011 it has risen to 96: 60 years of age and 36 of contributions. From 2013 it should have risen to 97: 61 years of age and 36 years of contributions.This last quota was never applied as it was repealed by a subsequent law);

- the "Fornero Reform" (Law n.214/2011) which constitutes the most significant recent pension reform in Italy. It introduced the 'flexible band' for workers who started contributing after '96 as they can retire between the ages of 63 and 70. Furthermore, it eliminated the quota system and eligibility criteria for old- age pension continued to increase and by 2050 the age requirement would become 69 years and 9 months for men and women for all types of workers. The most important innovation introduced by the reform, however, is the replacement of the seniority pension with the early retirement pension which no longer provides for a combination of age and years of contributions but exclusively for a contribution criterion. Under the new scheme are required 42 years for men in 2012, which will increase up to 46 years for men and 45 for women by 2050.[7]

Tables A.7 and A.8 summarize the retirement criteria made by the main reforms of the pension system over the years. To simplify the exposition of the different criteria, we have chosen to present two different tables to represent the pension criteria pre and post Fornero Reform. This division between early-retirement pension and old age pension is the one currently in force in Italy.

[7] In reality, those who are in the 100% contributory system can continue to retire by respecting a combination of contributions and minimum age but in this case the pension amount must not be less than 2.8 times the social allowance.

Bibliography

Abud, T., Kounidas, G., Martin, K. R., Werth, M., Cooper, K., and Myint, P. K. (2022). Determinants of healthy ageing: a systematic review of contemporary literature. *Aging clinical and experimental research*, 34(6):1215–1223.

Addabbo, T., Caiumi, A., and Maccagnan, A. (2012). The allocation of time within italian couples: exploring its unequal gender distribution and the effect of childcare services. *Annals of Economics and Statistics/ANNALES D'ÉCONOMIE ET DE STATISTIQUE*, pages 209–227.

Adjei, N. and Brand, T. (2018). Investigating the associations between productive housework activities, sleep hours and self-reported health among elderly men and women in western industrialised countries. *BMC Public Health*, 18:1–10.

Aguiar, M. and Hurst, E. (2005). Consumption versus expenditure. *Journal of political Economy*, 113(5):919–948.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Baker, L. and Silverstein, M. (2008). Depressive symptoms among grandparents raising grandchildren: The impact of participation in multiple roles. *Journal of Intergenerational Relationships*, 6(3):285–304.

Bank, W. (2020). *WOMEN, BUSINESS AND THE LAW - 2020*. World Bank, Washington, DC.

Bank, W. (2024). *Women, Business and the Law 2024*. The World Bank.

Barigozzi, F., Di Timoteo, C., and Monfardini, C. (2023). The gender gaps in time-use within italian households during 2002–2014. *Italian Economic Journal*, 9(3):1263–1296.

Battistin, E., Brugiavini, A., Rettore, E., and Weber, G. (2009). The retirement consumption puzzle: evidence from a regression discontinuity approach. *American Economic Review*, 99(5):2209–2226.

Battistin, E. and Rettore, E. (2008). Ineligibles and eligible non-participants as a double comparison group in regression-discontinuity designs. *Journal of Econometrics*, 142(2):715–730.

Bergeot, J. and Fontaine, R. (2020). The heterogeneous effect of retirement on informal care behavior. *Health economics*, 29(10):1101–1116.

Bertoni, M. and Brunello, G. (2021). Does a higher retirement age reduce youth employment? *Economic Policy*, 36(106):325–372.

Bettin, G., Giorgetti, I., and Staffolani, S. (2024). The impact of covid-19 lockdown on the gender gap in the italian labour market. *Review of Economics of the Household*, 22(1):1–33.

Bloemen, H. G., Pasqua, S., and Stancanelli, E. G. (2010). An empirical analysis of the time allocation of italian couples: are they responsive? *Review of Economics of the Household*, 8:345–369.

Boeri, T., Garibaldi, P., and Moen, E. (2017). Closing the retirement door and the lump of labor. *Consultabile in: http://sites. carloalberto. org/garibaldi/doc/papers/Lump_ of_ Labor_ 26may_ 2017. pdf.*

Bonsang, E. and Van Soest, A. (2020). Time devoted to home production and retirement in couples: a panel data analysis. *Labour Economics*, 65:101810.

Bostyn, R., Cherchye, L., De Rock, B., and Vermeulen, F. (2023). Time use, intrahousehold inequality, and individual welfare: Revealed preference analysis. In *Time Use in Economics*, pages 161–186. Emerald Publishing Limited.

Brini, E., Lenko, M., Scherer, S., and Vitali, A. (2021). Retraditionalisation? work patterns of families with children during the pandemic in italy. *Demographic research*, 45:957–972.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3):442–451.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.

Campaña, J. C., Gimenez-Nadal, J. I., and Velilla, J. (2023). Measuring gender gaps in time allocation in europe. *Social Indicators Research*, 165(2):519–553.

Carta, F. and De Philippis, M. (2024). The forward-looking effect of increasing the full retirement age. *The Economic Journal*, 134(657):165–192.

Carta, F., De Philippis, M., Rizzica, L., and Viviano, E. (2023). *Women, labour markets and economic growth*. Banca d'Italia.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2024a). *A practical introduction to regression discontinuity designs: Extensions*. Cambridge University Press.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2024b). *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Celidoni, M., Dal Bianco, C., and Weber, G. (2017). Retirement and cognitive decline. a longitudinal analysis using share data. *Journal of Health Economics*, 56:113–125.

Ciani, E. (2016). Retirement, pension eligibility and home production. *Labour Economics*, 38:106–120.

Commission, E. (2021). *Report on equality between women and men in the EU*.

Danielsbacka, M., Křenková, L., and Tanskanen, A. O. (2022). Grandparenting, health, and well-being: A systematic literature review. *European Journal of Ageing*, pages 1–28.

Del Boca, D., Oggero, N., Profeta, P., and Rossi, M. (2020). Women's and men's work, housework and childcare, before and during covid-19. *Review of Economics of the Household*, 18(4):1001–1017.

Del Boca, D., Oggero, N., Profeta, P., and Rossi, M. (2022). The impact of covid-19 on the gender division of housework and childcare: Evidence from two waves of the pandemic in italy. *IZA Journal of Labor Economics*, 11(1).

Dong, Y. and Lewbel, A. (2015). Identifying the effect of changing the policy threshold in regression discontinuity models. *Review of Economics and Statistics*, 97(5):1081–1092.

Eibich, P. and Siedler, T. (2020). Retirement, intergenerational time transfers, and fertility. *European Economic Review*, 124:103392.

Elefante, M., Hasan, T., Hyland, M., Martins, N. M. S., and Trumbic, T. (2023). Accelerating gender equality through reforming legal frameworks.

European Commission (2024). The 2024 pension adequacy report: current and future income adequacy in old age in the eu.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Heller-Sahlgren, G. (2017). Retirement blues. *Journal of Health Economics*, 54:66–78.

Hospido, L. and Zamarro, G. (2014). Retirement patterns of couples in europe. *IZA Journal of European Labor Studies*, 3:1–18.

Hurd, A. and Anderson, D. (2011). *Park and recreation professional's handbook with online resource*. Human Kinetics, Champaign, USA.

Hurd, M. D. and Rohwedder, S. (2003). The retirement-consumption puzzle: Anticipated and actual declines in spending at retirement.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of econometrics*, 142(2):615–635.

Istat (2016). *I tempi della vita quotidiana.*

Istat (2021). *Ritorni occupazionali dell'istruzione.*

Istat (2023). *INDICATORI DEMOGRAFICI - ANNO 2022*.

Lee, D. S. (2008). Randomized experiments from non-random selection in us house elections. *Journal of Econometrics*, 142(2):675–697.

Lee, D. S. and Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Mahne, K. and Huxhold, O. (2015). Grandparenthood and subjective well-being: Moderating effects of educational level. *The Journals of Gerontology: Series B*, 70(5):782–792.

Mangiavacchi, L., Piccoli, L., and Pieroni, L. (2021). Fathers matter: Intrahousehold responsibilities and children's wellbeing during the covid-19 lockdown in italy. *Economics & Human Biology*, 42:101016.

Marini, A. (2024). Updating the retirement-consumption puzzle in italy: who are the most affected?

Mazzonna, F. and Peracchi, F. (2012). Ageing, cognitive abilities and retirement. *European Economic Review*, 56(4):691–710.

Mazzonna, F. and Peracchi, F. (2017). Unhealthy retirement? *Journal of Human Resources*, 52(1):128–151.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

McCrary, J. and Royer, H. (2011). The effect of female education on fertility and infant health: evidence from school entry policies using exact date of birth. *American economic review*, 101(1):158–195.

Porter (2003). Estimation in the regression discontinuity model. unpublished manuscript. *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison, p 5-19*.

Randstad Research (2023). Gli inattivi: il lavoro non dichiarato, l'invecchiamento attivo, la sostenibilità al 2030 e al 2050. Available on line.

Schmitz, H. and Westphal, M. (2015). Short- and medium-term effects of informal care provision on female caregivers' health. *Journal of Health Economics*, 42:174–185.

SHARE-ERIC (2022a). Survey of Health, Ageing and Retirement in Europe (SHARE). Wave 8, release 8.0.0.

SHARE-ERIC (2022b). Survey of health, ageing and retirement in europe (share) wave 8. release version: 8.0.0.

Stancanelli, E. and Soest, A. V. (2012). Retirement and home production: A regression discontinuity approach. *American Economic Review*, 102(3):600–605.

Szinovacz, M. E. (2000). Changes in housework after retirement: A panel analysis. *Journal of Marriage and Family*, 62(1):78–92.

UN Women (2020). Progress of the world's women 2019-2020: families in a changing world, chapter 5. caring families, caring societies. Available on line.

United Nations Economic Commission for Europe, U. (2019). Active ageing index, analytical report.

van den Berg, B., Fiebig, D., and Hall, J. (2014). Well-being losses due to care-giving. *Journal of Health Economics*, 35:123–131.

World Health Organization (2002). Active ageing: A policy framework. Technical report, World Health Organization.

World Health Organization (2021). *Decade of healthy ageing: baseline report*. World Health Organization.

Zhang, K. and Zhang, W. (2019). Leisure activities and healthy aging. In Gu, D. and Dupre, M., editors, *Encyclopedia of Gerontology and Population Aging*. Springer, Cham, Switzerland.

Table A.7: Pension Requirements Pre Fornero for Men and Women

Year	Men						Women					
	Seniority			Old Age			Seniority			Old Age		
	Private	Public	Self-Employed	Private	Public	Self-Employed	Private	Public	Self-Employed	Private	Public	Self-Employed
1973 - 1993	35	35	35	60	65	60	35	35	35	55	60	55
1994	35	35	35	61	65	61	35	35	35	56	60	56
1995	35	35	35	61	65	61	35	35	35	56	60	56
1996	52 and 35 (36)	52 and 35 (36)	56 and 35 (40)	62	65	62	52 and 35 (36)	52 and 35 (36)	57 and 35 (40)	57	60	57
1997	52 and 35 (36)	52 and 35 (36)	56 and 35 (40)	63	65	63	52 and 35 (36)	52 and 35 (36)	56 and 35 (40)	57	60	57
1998	54 and 35 (36)	53 and 35 (36)	57 and 35 (40)	63	65	63	54 and 36 (36)	53 and 36 (36)	56 and 36 (40)	58	60	58
1999	55 and 35 (37)	53 and 35 (37)	57 and 35 (40)	64	65	64	55 and 36 (37)	53 and 36 (37)	57 and 36 (40)	59	60	59
2000	55 and 35 (37)	54 and 35 (37)	57 and 35 (40)	65	65	65	55 and 36 (37)	54 and 36 (37)	57 and 36 (40)	60	60	60
2001	56 and 35 (37)	55 and 35 (37)	58 and 35 (40)	65	65	65	56 and 35 (37)	55 and 35 (37)	58 and 35 (40)	60	60	60
2002	57 and 35 (37)	55 and 35 (37)	58 and 35 (40)	65	65	65	57 and 35 (37)	55 and 35 (37)	58 and 35 (40)	60	60	60
2003	57 and 35 (37)	56 and 35 (37)	58 and 35 (40)	65	65	65	57 and 35 (37)	56 and 35 (37)	58 and 35 (40)	60	60	60
2004	57 and 35 (38)	57 and 35 (38)	58 and 35 (40)	65	65	65	57 and 35 (38)	57 and 35 (38)	58 and 35 (40)	60	60	60
2005	57 and 35 (38)	57 and 35 (38)	58 and 35 (40)	65	65	65	57 and 35 (38)	57 and 35 (38)	58 and 35 (40)	60	60	60
2006	57 and 35 (39)	57 and 35 (39)	58 and 35 (40)	65	65	65	57 and 35 (39)	57 and 35 (39)	58 and 35 (40)	60	60	60
2007	57 and 35 (39)	57 and 35 (39)	58 and 35 (40)	65	65	65	57 and 35 (39)	57 and 35 (39)	58 and 35 (40)	60	60	60
2008	58 and 35 (40)	58 and 35 (40)	59 and 35 (40)	65	65	65	58 and 35 (40)	58 and 35 (40)	58 and 35 (40)	60	60	60
2009	59 and 35 (40)	59 and 35 (40)	60 and 35 (40)	65	65	65	59 and 35 (40)	59 and 35 (40)	59 and 35 (40)	60	60	60
2010	59 and 35 (40)	59 and 35 (40)	60 and 35 (40)	65	65	65	59 and 35 (40)	59 and 35 (40)	60 and 35 (40)	60	60	60
2011	60 and 35 (40)	60 and 35 (40)	61 and 35 (40)	65	65	65	60 and 35 (40)	60 and 35 (40)	61 and 35 (40)	60	60	60

Note: the table shows the two methods through which eligibility for seniority pension can be achieved: the age requirement along with the minimum required contributions, or the fulfillment of the contribution requirement alone (indicated in parentheses).
With regard to the old-age pension, only the age requirement is indicated; however, in determining the eligibility age, the necessary minimum contribution has also been considered. Additionally, certain exceptions for reaching the old-age pension depend on when the worker started working (before or after 1996). Although these exceptions are not explicitly indicated in the table, they have been taken into account in the calculation of the age at eligibility.
It should also be noted that the Amato Reform, initiated in 1992, led to changes in the criteria every one and a half years instead of each calendar year, which has resulted in certain inaccuracies in the table.

Table A.8: Pension Requirements Post Fornero for Men and Women

Year	Men						Women					
	Early retirement			Old Age			Early retirement			Old Age		
	Private	Public	Self-Employed	Private	Public	Self-Employed	Private	Public	Self-Employed	Private	Public	Self-Employed
2012	42.1 (63)	42.1 (63)	42.1 (63)	66	66	66	41.1 (63)	41.1 (63)	41.1 (63)	62	66	63.6
2013	42.5 (63.3)	42.5 (63.3)	42.5 (63.3)	66.3	66.3	66.3	41.5 (63.3)	41.5 (63.3)	41.5 (63.3)	62.3	66.3	63.9
2014	42.6 (63.3)	42.6 (63.3)	42.6 (63.3)	66.3	66.3	66.3	41.6 (63.3)	41.6 (63.3)	41.6 (63.3)	63.9	66.3	64.9
2015	42.6 (63.3)	42.6 (63.3)	42.6 (63.3)	66.3	66.3	66.3	41.6 (63.3)	41.6 (63.3)	41.6 (63.3)	63.9	66.3	64.9
2016	42.10 (63.7)	42.10 (63.7)	42.10 (63.7)	66.7	66.7	66.7	41.10 (63.7)	41.10 (63.7)	41.10 (63.7)	65.7	66.7	66.1
2017	42.10 (63.7)	42.10 (63.7)	42.10 (63.7)	66.7	66.7	66.7	41.10 (63.7)	41.10 (63.7)	41.10 (63.7)	65.7	66.7	66.1
2018	42.10 (63.7)	42.10 (63.7)	42.10 (63.7)	66.7	66.7	66.7	41.10 (63.7)	41.10 (63.7)	41.10 (63.7)	66.7	66.7	66.7
2019	42.10(64)	42.10(64)	42.10(64)	67	67	67	41.10 (64)	41.10 (64)	41.10 (64)	67	67	67
2020	42.10(64)	42.10(64)	42.10(64)	67	67	67	41.10 (64)	41.10 (64)	41.10 (64)	67	67	67

Note: the table shows the two methods through which eligibility for early retirement pension can be achieved: the fulfillment of the contribution requirement alone, or the age requirement (indicated in parentheses) plus twenty years of contributions. However, this last possibility is available only for workers who started working after 1995.
With regard to the old-age pension, only the age requirement is indicated; however, in determining the eligibility age, the necessary minimum contribution has also been considered. Additionally, certain exceptions for reaching the old-age pension depend on when the worker started working (before or after 1996). Although these exceptions are not explicitly indicated in the table, they have been taken into account in the calculation of the age at eligibility.

Chapter two

`rdlasso`: Regression Discontinuity with High-Dimensional Data

Marianna Nitti* Marco Ventura*

*Sapienza University of Rome, Rome, Italy

Abstract

We present a command, `rdlasso`, which allows the inclusion of high-dimensional covariates in Regression Discontinuity Design (RDD) settings. This command is based on the paper "Inference in Regression Discontinuity Designs with High-Dimensional Covariates" by Kreiss and Rothe (2023). The command automates covariate selection via Lasso-based procedures, supports both sharp and fuzzy settings and integrates seamlessly with `rdrobust` for bandwidth selection and inference. The implementation relies on a Python backend to perform the high-dimensional model selection, making the methodology accessible to Stata users and computationally feasible for applied researcher.

Keywords: Regression Discontinuity Design, Lasso, high-dimensional data, machine learning

JEL Codes: C10, C20, C21, C26

1 Introduction

Including many predetermined covariates can improve the accuracy of treatment effect estimates in Regression Discontinuity Designs (RDD). To leverage this clear advantage Kreiss and Rothe (2023) set out an estimator consisting of two steps: in the first step, a localised Lasso-type procedure is used to select a reduced set of covariates; in the second step, the covariates selected are used to estimate the treatment effect following the standard local linear estimation methodology. From a theoretical point of view, under an approximate sparsity condition, the estimator is conceptually similar to the usual local linear estimator in terms of both bias and variance and benefits from the asymptotic properties of localized estimators under broader classes of machine learning (ML) methods.

Python provides powerful frameworks for executing ML algorithms, among which the most widely used is undoubtedly scikit-learn, supporting the fitting of numerous ML models. Starting from Stata 16, an application programming interface (API)[1] has made it possible to integrate Stata with Python. Although this integration has expanded Stata's computational capabilities, the software still lacks a built-in command that directly implements the method proposed by Kreiss and Rothe (2023). To address this gap, we developed the `rdlasso` package, which allows researchers to implement RDD with high-dimensional covariates in Stata, offering the possibility to handle both sharp and fuzzy cases through a simple option. The `rdlasso` command combines Python-based ML tools with Stata's consolidated local linear estimation procedures for the estimation of treatment effects, following the framework of `rdrobust` by Calonico et al. (2017). Despite Stata already includes built-in Lasso implementations, integration with Python provides a more general and extensible framework, consistent with the theoretical approach developed by Kreiss and Rothe (2023).

The advantage offered by `rdlasso` is threefold. First, it allows Stata users to apply ML methods in a RDD setup characterised by high-dimensional covariates without the need to learn other software environments (specifically R or Python). Second, the command generates variables that can be used in subsequent Stata codes in the same do-file. Third, `rdlasso` is fully compatible with `rdrobust`: the Python backend returns the optimal bandwidths, which are then passed to `rdrobust` for estimation and inference. This ensures that users can rely on the familiar `rdrobust` framework for robust inference and easily perform sensitivity or robustness analyses using the same bandwidths. The command is itself an interesting case of Stata/Python integration that can accommodate both sharp and fuzzy designs through a simple option.

The remainder of the article proceeds as follows. Section 2 briefly recalls the econometrics underpinning the estimator being implemented; Section 3 presents the syntax of the `rdlasso` command; Section 4 shows an application of the command on a case study, and, finally Section 5 concludes the article.

[1] An application programming interface (API) is a set of tools and routines that enables two software environments to communicate and exchange data efficiently. In Stata, this functionality is implemented through the Stata Function Interface (sfi) module, which provides Python access to Stata's core features such as datasets, macros, matrices, and scalars.

2 RDD with high-dimensional covariates

In many empirical applications, researchers include covariates to improve the precision of the estimates. On the one hand, the inclusion of covariates may improve precision; on the other hand, given the limited number of observations in the bandwidth, the addition of covariates can generate distorted inferences (Calonico et al. (2014); Calonico et al. (2019)). This issue becomes even more pronounced in high-dimensional settings, where the number of covariates exceeds the number of observations, a common occurrence in modern empirical work involving rich datasets or extensive transformations of covariates.

To address this problem Kreiss and Rothe (2023) propose a two-step approach. In the first step, they select a small subset of covariates by adding an ℓ_1 penalty (see Tibshirani, 1996) to the local least squares problem and retain those variables with non-zero estimated coefficients. By construction, the selected covariates are strongly correlated with the outcome and therefore have the greatest potential to absorb part of its variance. In the second step, they compute a standard linear adjustment estimator using only the selected covariates.
The authors show that the resulting "post-Lasso" estimator is asymptotically normal under an "approximate sparsity" condition, with asymptotic bias and variance similar to those obtained in low-dimensional settings. Approximate sparsity refers to situations in which only a small number of covariates are truly relevant for empirical analysis, meaning that including additional covariates would not substantially improve the accuracy of the estimates.
In this 2-steps procedure standard bandwidth selection and inference methods can be applied to the selected variables, making the approach easy to implement.

Despite some differences, the estimator shares important features with the post-Lasso approach proposed by Belloni et al. (2014) for estimating treatment effects under unconfoundedness with high-dimensional controls. In the RDD setting, covariate selection is not required for consistency, as covariates serve only to improve efficiency, not to address selection bias. Therefore, Lasso in the RDD context is performed in a single selection step targeting variables predictive of the outcome, in contrast to the double selection procedure as in Belloni et al. (2014), which also includes variables related to treatment. Moreover, the variable selection step is *local* in the sense that it is computed within the optimal bandwidth, rather than through the standard Lasso. Unlike the unconfoundedness framework, RDD cannot rely on (conditionally) random treatment assignment; instead, it exploits continuity assumptions. As a result, Kreiss and Rothe (2023) derive *ad hoc* conditions that differ from that of Belloni et al. (2014).

2.1 The problem from a formal point of view. The sharp case

In a standard RDD set up we observe a random sample $\{(Y_i, X_i, W_i)\}_{i=1,\dots,n}$ where $Y_i \in \mathbb{R}$ is the outcome variable, $X_i \in \mathbb{R}$ is the running variable (or score), and $W_i \in \mathbb{R}^p$ is a vector of pre-treatment covariates. The focus is on high-dimensional settings, where the number of covariates p may be large relative to (or even exceed) the sample size n. For the time being, treatment is assigned if and only if the running variable exceeds a known threshold, normalized to 0: $T_i = \mathbb{I}(X_i \geq 0)$. Each unit has potential outcomes $Y_i(0)$ and $Y_i(1)$, and the observed outcome is

$$Y_i = Y_i(T_i)$$

The parameter of interest is the *average treatment effect (ATE) at the threshold*:

$$\tau_Y = \mathbb{E}[Y_i(1) - Y_i(0) \mid X_i = 0]$$

Under the continuity assumption of the conditional expectations $\mathbb{E}[Y_i(t) \mid X_i = x]$ for $t \in \{0, 1\}$, this quantity is identified via the discontinuity in the conditional mean of Y_i:

$$\tau_Y = \lim_{x \downarrow 0} \mathbb{E}[Y_i \mid X_i = x] - \lim_{x \uparrow 0} \mathbb{E}[Y_i \mid X_i = x]$$

A common estimator for τ_Y is the local linear regression estimator, which fits separate linear models on either side of the threshold. The *baseline estimator* is:

$$\hat{\tau}_{h,\text{Base}} = e_2^\top \arg\min_{\theta \in \mathbb{R}^4} \sum_{i=1}^{n} K_h(X_i) \left[Y_i - V_i^\top \theta \right]^2 \tag{2.1}$$

where:

- $K_h(x) = \frac{1}{h} K\left(\frac{x}{h}\right)$ is a kernel function with bandwidth h,
- $V_i = (1, T_i, X_i/h, T_i X_i/h)^\top$,
- $e_2 = (0, 1, 0, 0)^\top$ selects the coefficient of interest, i.e. the treatment effect given by the coefficient of T_i.

Researchers often augment eq. (2.1) with covariates, giving rise to the covariate-adjusted linear estimator (Calonico et al., 2019):

$$\hat{\tau}_{h,\text{CCFT}} = e_2^\top \arg\min_{(\theta,\gamma) \in \mathbb{R}^{4+p_n}} \sum_{i=1}^{n} K_h(X_i) \left[Y_i - V_i^\top \theta - W_i^\top \gamma \right]^2 \tag{2.2}$$

This estimator is consistent under standard identification conditions typical of RDD settings; see Cattaneo et al. (2024) for a recent extended overview of the literature.[2] However, in high-dimensional contexts, this estimator becomes undefined, particularly when the number of covariates exceeds the number of effective observations. Moreover, it might not provide adequate descriptions of the estimator's finite sample properties due to overfitting, even with moderate dimensionality.

Kreiss and Rothe (2023) propose the following two-step procedure to extend covariate-adjusted estimation to high-dimensional settings.

[2]Notice that the formulation in eq. (2.1) is also consistent with Regression Kink Design, RKD, by considering the coefficient of $T_i X_i$ which accounts for differences in slopes.

Step 1: Local Lasso Selection. Given a preliminary bandwidth b and a penalty parameter λ, solve the following weighted Lasso problem:

$$(\hat{\theta}_n, \hat{\gamma}_n) = \arg \min_{(\theta,\gamma)\in\mathbb{R}^{4+p_n}} \sum_{i=1}^{n} K_b(X_i) \left[Y_i - V_i^\top \theta - (W_i - \hat{\mu}_{W,n})^\top \gamma\right]^2 + \lambda \sum_{k=1}^{p_n} \hat{w}_{n,k} |\gamma_k| \tag{2.3}$$

where

$$\hat{\mu}_{W,n} = \frac{1}{n} \sum_{i=1}^{n} W_i K_b(X_i)$$

is the local sample mean of the covariates and

$$\hat{w}_{n,k}^2 = \frac{1}{n} \sum_{i=1}^{n} K_b(X_i) \left(W_i^{(k)} - \hat{\mu}_{W,n}^{(k)}\right)^2$$

is the local variance for the k-th covariate. The covariates are not explicitly standardized, but in eq. (2.3) we subtract the local mean of the covariates and multiply the Lasso penalty by a weight, $\hat{w}_{n,k}$, which equals the local standard deviation of the individual covariates, both actions together have the same effect as providing standardized data. In this way, variables which change only a little receive a lower penalty which is desirable because the coefficient needs to be larger in order to have an effect in the least squares criterion. Put differently, standardising the covariates allows the penalty parameter λ to be reasonably tuned to all coefficients simultaneously. The selected covariate set is defined as $\hat{J}_n = \left\{k \in \{1, \ldots, p_n\} : \hat{\gamma}_n^{(k)} \neq 0\right\}$

Step 2: Post-Lasso Estimation. Using a final bandwidth h, compute the restricted post-Lasso estimator:

$$\hat{\tau}_h(\hat{J}_n) = e_2^\top \arg \min_{(\theta,\gamma)\in\mathbb{R}^{4+|\hat{J}_n|}} \sum_{i=1}^{n} K_h(X_i) \left[Y_i - V_i^\top \theta - W_i(\hat{J}_n)^\top \gamma\right]^2 \tag{2.4}$$

The asymptotic distribution of $\hat{\tau}_h(\hat{J}_n)$ does not depend on the preliminary bandwidth b or the penalty λ, although these parameters must be chosen in practice. In fact, the preliminary bandwidth b can be selected using methods designed for baseline RD estimators without covariates, for example Imbens and Kalyanaraman (2012) Calonico et al. (2017), Armstrong and Kolesár (2018).

The penalty parameter λ can be chosen by cross-validation or via a plug-in procedure as in Belloni et al. (2014).
In analogy with the case of b the final bandwidth h may be selected using standard techniques appropriate for linear adjustment with low-dimensional covariates (Calonico et al., 2017; Armstrong and Kolesár, 2018).

The authors formally establish the asymptotic normality of the estimator plus a number of desirable properties that can be summarised as follows:

- **Asymptotic Bias:** the covariate adjustment component of the bias disappears under some regularity conditions (see Kreiss and Rothe, 2023, Theorem 1)

- **Asymptotic Variance:** the estimator achieves optimal asymptotic variance among all linear adjustment estimators using a moderately sized subset of covariates, even when exact sparsity does not hold.

- **Double Selection:** unlike double selection procedures common in estimation of treatment effect under unconfoundedness (e.g., Belloni et al., 2013), the incorporation of predictive covariates of treatment assignment is unnecessary in settings of RDD, as treatment is locally as good as random assigned.

- **Role of the Lasso:** the methodology does not require exact recovery of the true covariate set; it suffices that the model selection procedure generates residuals similar to those generated by the true set. It follows that alternative selection methods that meet the same condition could be used.

Overall, the results provide strong theoretical support for the post-Lasso approach in RD settings with high-dimensional covariates, highlighting its flexibility, robustness, and efficiency.

2.2 The fuzzy case

The estimator can be easily extended to a fuzzy setup. In fuzzy RDD, not all units necessarily comply with the treatment assignment. Therefore, in contrast with sharp RDD, the conditional treatment probability at the cutoff does not jump from zero to one, but rather takes a value in between. In this case, one must distinguish between eligibility, Z_i, and treatment T_i. Put simply, the extension from sharp to fuzzy, amounts to moving from a linear regression model to an instrumental variable approach, where Z_i is the instrument. In this case, the causal effect of T_i on Y_i is defined as the Local Average Treatment Effect (LATE) concerning only the sub population of compliers and is estimated as the ratio of the jump in the outcome, Y_i, to the jump in the probability of being treated, at the cutoff.
In more details, the jump in the outcome is estimated by replacing T with Z in Step 1 and 2 before, namely in eq. (2.3) and eq. (2.4), this stage will represent the Lasso-type equivalent of the so called reduced form equation. Similarly, the jump in the probability of being treated is estimated by replacing Y with T and T with Z in Step 1 and 2 before, namely in eq. (2.3) and eq. 2.4, this stage will represent the first-stage Lasso-type equation. It follows that the coefficients of interest will become the coefficient of Z in both regressions, i.e. θ_2 and π_2, respectively. Hence, the fuzzy parameter can be obtained by repeating Step 1 and 2 twice.
More formally, defining by $F_i = (1, Z_i, X_i/h, Z_iX_i/h)^\top$ θ_2 is obtained as:

Step 1: reduced form Lasso-type

$$(\hat{\theta}_n, \hat{\gamma}_n) = \arg \min_{(\theta,\gamma)\in\mathbb{R}^{4+p_n}} \hat{\tau}_h(\hat{J}_n) = \sum_{i=1}^{n} K_b(X_i) \left[Y_i - F_i^\top\theta - (W_i - \hat{\mu}_{W,n})^\top\gamma\right]^2 + \lambda \sum_{k=1}^{p_n} \hat{w}_{n,k}|\gamma_k| \tag{2.5}$$

Step 2: reduced form Lasso-type

$$\hat{\theta}_2 = e_2^\top \arg \min_{(\pi,\gamma)\in\mathbb{R}^{4+|\hat{J}_n|}} \sum_{i=1}^{n} K_h(X_i)\left[Y_i - F_i^\top \theta - W_i(\hat{J}_n)^\top \gamma\right]^2 \tag{2.6}$$

Similarly, π_2 is obtained as:

Step 1: first stage Lasso-type

$$(\hat{\pi}_n, \hat{\gamma}_n) = \arg \min_{(\theta,\gamma)\in\mathbb{R}^{4+p_n}} \hat{\tau}_h(\hat{J}_n) = \sum_{i=1}^{n} K_b(X_i)\left[T_i - F_i^\top \pi - (W_i - \hat{\mu}_{W,n})^\top \gamma\right]^2 + \lambda \sum_{k=1}^{p_n} \hat{w}_{n,k}|\gamma_k| \tag{2.7}$$

Step 2: first stage Lasso-type

$$\hat{\pi}_2 = e_2^\top \arg \min_{(\pi,\gamma)\in\mathbb{R}^{4+|\hat{J}_n|}} \sum_{i=1}^{n} K_h(X_i)\left[T_i - F_i^\top \pi - W_i(\hat{J}_n)^\top \gamma\right]^2 \tag{2.8}$$

The fuzzy RDD treatment effect is then $\tau_{FRD} = \theta_2/\pi_2$, which is the ratio of two sharp RD estimands, with the standard error of τ_{FRD} computed *via* delta method (see, among others Cox, 1990).

3 The command

3.1 Command prerequisites

The `rdlasso` command requires Stata 17.0 or higher and a working installation of Python 3.8. The command implements the methodology of Kreiss and Rothe (2023) using custom Python routines for high-dimensional model selection, integrated into Stata via the Stata Function Interface (`sfi`). The Anaconda distribution can be used but is not mandatory. The command requires the installation of several Python packages and their dependencies. In particular, the following list reports the required software components and the corresponding minimum versions used for development and testing:

- `rdrobust` (Python package);
- numpy (version 1.17.4 or higher);
- pandas (version 1.0.0 or higher);
- scikit-learn (version 0.22 or higher);
- SciPy (version 1.3.3 or higher);
- statsmodels (version 0.10.2 or higher).

In addition, the implementation makes use of Python's native `os`, `sys`, and `warnings` modules.

A proper functioning of the Stata Function Interface (`sfi`) Python module shipped with the installed Stata version and flavor is required, as it provides the bidirectional communication between Stata and Python that underlies the execution of the command.

On the Stata side, the command requires the installation of `rdrobust` (Calonico et al., 2017). If not already installed, it can be dowloaded from Stata by typing `ssc install rdrobust`.

Detailed step-by-step instructions for configuring Python and installing the required packages are provided in the Appendix.

3.2 The `rdlasso` command

Syntax

The syntax of the command is as follows:

`rdlasso` *varlist(min=3)* [*if*] [*in*] [,`t` (*varname max=1*) `z` (*varname max=1*) `c` (*real*) `fuzzy` (*string*) `level` (*real*) `b` (*real*) `bfactor` (*real*) `h` (*real*) `kernel` (*string*)]

varlist is a list of numerical variables, it must include in order: the outcome variable (Y), the running variable (X), the list of covariates (D) to be used for high-dimensional adjustment.

Treatment and instrument variables are not included in *varlist*. In the fuzzy RD case, the treatment variable (T) and the instrument (Z) must be specified through the options `t()` and `z()`, respectively. In the sharp RD case, the treatment variable may be specified or omitted; if omitted is automatically constructed as an indicator for being above the cutoff.

Options

`c` (*real*) specifies the cutoff value of the running variable (default is 0).

`fuzzy` (*string*) activates fuzzy design logic (default is "off"), `fuzzy ("on")` is required for fuzzy designs.

`t` (*varname max=1*) specifies the treatment variable, `t ()` is required for fuzzy designs.

`z` (*varname max=1*) specifies the instrumental variable, `z ()` is required for fuzzy designs.

`level` (*real*) indicates the confidence level for inference, expressed as a percentage. The default value is 95.

`b` (*real*) specifies the bandwidth used for the model selection step. If not specified, the bandwidth is selected using `rdrobust` without covariates.

`bfactor` (*real*) indicates the scaling factor applied to the selected bandwidth for the model selection (default value is 1).

`h` (*real*) specifies the bandwidth used for the final RD estimation. If not specified, it is selected using `rdrobust` with the covariates selected in the model selection step.

`kernel` (*string*) specifies the kernel function to be used for both model selection and estimation. Possible choices are triangular (default), epanechnikov, uniform.

Stored results

In the *sharp* case, `rdlasso` internally calls `rdrobust` for estimation, so that all results are stored in `e()` and mirror those of the original `rdrobust` command, as documented in Calonico et al. (2017). In this case, the point estimate and the robust standard error can be retrieved directly from `e(tau_bc)` and `e(se_tau_rb)`, respectively.

In the *fuzzy* case, instead, `rdlasso` performs two separate sharp regressions—one for the first stage and one for the reduced form—and computes the fuzzy estimator as the ratio between the two bias-corrected point estimates. The local average treatment effect and its inference are therefore obtained from the following scalars stored in Stata:

e(tau_fuzzy) fuzzy RD point estimate

e(se_fuzzy) robust standard error

e(z_fuzzy) z-statistic

e(p_fuzzy) p-value

e(ci_low) lower bound of the 95% confidence interval

e(ci_high) upper bound of the 95% confidence interval

Remarks

All variables supplied to `rdlasso` must be numeric. The command does not automatically handle factor variables or categorical covariates. If categorical variables are to be included, the user is responsible for transforming them into appropriate numerical representations (e.g., dummy variables) prior to estimation.

Moreover, the command does not internally handle missing values. Observations with missing values in any of the variables supplied to *varlist* or to the options `t()` and `z()` should be addressed by the user before running the command.

Finally, variable names must comply with Stata naming conventions and be uniquely identifiable. Potential issues related to ambiguous or duplicated variable names are particularly likely to arise in high-dimensional settings, where users often generate large sets of transformed or interaction-based covariates programmatically. To avoid errors during data transfer between Stata and the Python backend, users should therefore pay special attention to consistent and unambiguous naming of covariates.

4 Applications

4.1 The sharp case

For the sharp RD case we use the illustrative example from Cattaneo et al. (2019), based on the study by Meyersson (2014). This study analyzes the effect of Islamic political representation in Turkish municipal elections on women's education achievement. The observational units are municipalities, and the running variable is represented by the Islamic party's margin of victory in the 1994 Turkish mayoral elections. The treatment is the electoral victory of Islamic parties, and the cutoff is zero. Specifically, municipalities elect an Islamic mayor when the Islamic vote margin is positive; otherwise, they elect a secular mayor. The variables of interest are reported below:

- Y: share of the 15–20-year cohort who, in the 2000 census, were recorded to have completed high school;
- X: the difference in vote share between the largest Islamic party and the largest secular party, resulting in a cutoff at zero;
- T: is the treatment dummy with $T_i = \mathbf{1}(X_i \geq 0)$.
- W: baseline set of covariates containing Islamic vote percentage in 1994 (*vshr_ islam1994*), the number of parties receiving votes in 1994 (*partycount*), the logarithm of the population in 1994 (*lpop1994*), a district center indicator (*merkezi*), a province center indicator (*merkezp*), a sub-metro center indicator (*subbuyuk*), and a metro center indicator (*buyuk*).

```
. help rdlasso
```

By default, `rdlasso` sets the value of `cutoff` to zero, the value of `kernel` to `Triangular`, the value of `level` to 0.95, and the value of `bfactor` to 1.0. All these settings can be modified using their respective options. Regarding the choice of `b`, Kreiss and Rothe (2023) suggest using the bandwidth selection method designed for settings without covariates, such as those proposed by Imbens and Kalyanaraman (2012) and Calonico et al. (2014). Although the user can manually choose the final bandwidth `h`, in the command we select it by applying the standard methodology used in the low-dimensional case with covariates Calonico et al. (2017).

The choice of λ is based on the plug-in procedure proposed by Belloni et al. (2014), which was originally developed for non-localized Lasso estimators and is therefore adapted here to the RD context. Moreover, adjustments are introduced to the plug-in procedure to account for the presence of kernel weights.

To accommodate the high-dimensional setting, in addition to the baseline set of covariates used in Cattaneo et al. (2019), we include second-order polynomial terms and pairwise interactions for all continuous variables, as well as dummy variables for categorical covariates. This results in a total of 170 covariates for $2,629$ observations. Subsequently, the RD parameter, its standard error, and confidence intervals were calculated using `rdlasso`.

Under standard conditions (e.g., covariate balance at the cutoff), we expect the variance of the covariate-adjusted estimator to be not higher than that of the baseline estimator.

The `rdlasso` command is executed as follows:

```
. rdlasso Y X ageshr19-prov_num_Zonguldak
```

The set of selected covariates for the outcome variable Y is identified through the corresponding column indices:

```
 Selected covariates (Y):  ageshr19_2 ageshr60_merkezi hischshr1520m_2 hischshr1520m_partycount
-----------------------------------------------------------
 Bandwidth h (from Python): 10.99236114026833
-----------------------------------------------------------
```

Figure 1: rdlasso - selected covariates

```
Covariate-adjusted Sharp RD estimates using local polynomial regression.

     Cutoff c = 0 | Left of c  Right of c            Number of obs =       2629
------------------+----------------------            BW type       =     Manual
    Number of obs |      2314         315            Kernel        = Triangular
Eff. Number of obs|       342         206            VCE method    =         NN
    Order est. (p)|         1           1
    Order bias (q)|         2           2
       BW est. (h)|    10.992      10.992
      BW bias (b) |    10.992      10.992
        rho (h/b) |     1.000       1.000

Outcome: Y. Running variable: X.
--------------------------------------------------------------------------------
                  | Point      | Robust Inference
                  | Estimate   | z-stat        P>|z|     [95% Conf. Interval]
------------------+------------+------------------------------------------------
        RD Effect |  2.1416    | 2.0743        0.038      .143702      5.07158
--------------------------------------------------------------------------------
Covariate-adjusted estimates. Additional covariates included: 4
```

Figure 2: rdlasso - covariate-adjusted RD estimates

The `rdlasso` output format, Figure 2, mirrors that of `rdrobust` in the sharp RD case. In the upper panel, information about the total number of observations, the bandwidths used, the kernel type, and other details are reported. In the lower part, the estimation results are displayed. The RD point estimate is reported in the bottom panel, together with the robust standard error, the standardized test statistic, the corresponding p-value, and the robust bias-corrected 95% confidence interval.

Our method produces a point estimate of 2.1416 percentage points, which is quite close to the one obtained using `rdrobust` without covariates (3.0195) and with a baseline set of covariates (3.108).[3] As excpected, `rdlasso` is remarkably more efficient, with a standard error of 0.038 against 1.488 and 1.700 from `rdrobust` with and without covariates, respectively. Compared to both previous approaches, this shows that a small, carefully selected set of covariates can indeed improve the precision of the estimates. As an obvious consequence, the corresponding confidence interval is also shorter by a similar factor. The plug-in procedure selects only four of the 170 covariates, corresponding to: *ageshr19_2, ageshr60_merkezi, hischshr1520m_2, hischshr1520m_partycount*, where the first is the square of the percentage of the population under 19 in 2000, the second is the percentage of the population over 60 in 2000 interacted with the district center, the third is the square of the percentage of men aged 15-20 with a high school diploma, and the last is the square of the percentage of men aged 15-20 with a high school diploma interacted with the number of parties that received votes in 1994.

[3]The results of Cattaneo et al. (2019), are provided in Appendix, see Table A.1

4.2 The fuzzy case

In this section, we present an application of `rdlasso` to a fuzzy RD case. We rely on the teaching example from Cattaneo et al. (2024), based on the study by Londoño-Vélez et al. (2020). This study focuses on the effects of the Ser Pilo Paga (SPP) scholarship programme in Colombia. The programme aimed to finance the full tuition of a four- or five-year undergraduate degree at a government-certified Higher Education Institution (HEI) that met minimum quality standards. To be eligible for the scholarship, students had to meet two requirements: achieve high academic performance and come from economically disadvantaged families, as determined by a survey-based wealth index (SISBEN). The variables of interest are defined as follows:

- Y: enrollment in a high-quality higher education institution (a binary indicator);
- X: difference between the student's SISBEN wealth index and its corresponding eligibility threshold (the cutoff is therefore normalized to zero);
- Z: eligibility dummy equal to 1 for eligible students;
- T: treatment dummy equal to 1 for students receiving the scholarship.
- W: baseline set of predetermined covariates including an indicator equal to one if the student identifies as female (*icfes_female*), the student's age (*icfes_ age*), an indicator equal to one if the student identifies as an ethnic minority (*icfes_ urm*), the residential stratum of the student's household (*icfes_ stratum*), and the student's family size (*icfes_famsize*).

All general considerations about syntax, default options, and bandwidth selection discussed in the sharp RD case still apply. However, the fuzzy design introduces some important differences.

Specifically, three options become mandatory: `fuzzy` to activate the fuzzy design logic, which must be set to `"on"`; `t` to specify the treatment variable; `z` to specify the instrumental variable.

In the fuzzy case, two regressions are carried out: a first regression in which the treatment variable T is regressed on the instrument Z (first stage), and a second in which the outcome variable Y is regressed on the same instrument Z (reduced form). Because the Lasso-based selection algorithm identifies the relevant covariate sets independently for each equation, the set of variables selected in the two regressions may differ. It follows that `rdlasso` calls `rdrobust` twice, using in each case the covariates selected by Lasso in Python. Finally, the coefficient of interest is obtained as the ratio between the reduced form coefficient and the first stage coefficient, while the standard errors are calculated using the Delta Method, whence the corresponding p-values and confidence intervals.

We run the rdlasso command as follows:

```
. rdlasso Y X icfes_female-icfes_famsize_4, t(T) z(Z) fuzzy(on)
```

The set of covariates selected in the two steps is shown below:

```
Selected covariates (Y):  icfes_age_4
Selected covariates (T):  icfes_age_3
-----------------------------------------------------------------
Bandwidth h_Y (from Python): 8.741355972886943
Bandwidth h_T (from Python): 18.73221556833528
```

Figure 3: rdlasso - selected covariates - fuzzy case

```
Covariate-adjusted Sharp RD estimates using local polynomial regression.

      Cutoff c = 0 | Left of c  Right of c            Number of obs =      23132
-------------------+----------------------            BW type       =     Manual
     Number of obs |      7709       15423            Kernel        = Triangular
Eff. Number of obs |      6669        7542            VCE method    =         NN
    Order est. (p) |         1           1
    Order bias (q) |         2           2
       BW est. (h) |    18.732      18.732
      BW bias (b)  |    18.732      18.732
         rho (h/b) |     1.000       1.000

Outcome: T. Running variable: X.
--------------------------------------------------------------------------------
                   | Point         | Robust Inference
                   | Estimate      | z-stat       P>|z|     [95% Conf. Interval]
-------------------+---------------+--------------------------------------------
         RD Effect |  .62631       | 35.5242      0.000      .59113      .660167
--------------------------------------------------------------------------------
Covariate-adjusted estimates. Additional covariates included: 1
```

Figure 4: rdlasso - First stage

The output of **rdlasso** reports the results of the two stages separately. In our case, the first stage is in Figure 4 and the reduced form in Figure 5, whereas the final result is in Figure 6. The substantial difference between **rdlasso** and **rdrobust**'s results lies in the selection of covariates, since the **rdrobust** command includes five covariates (see Table A.1 in the appendix) while **rdlasso** selects only one, *icfes_ age_ 3*, corresponding to the cube of age.

Covariate-adjusted Sharp RD estimates using local polynomial regression.

Cutoff c = 0	Left of c	Right of c
Number of obs	7709	15423
Eff. Number of obs	3771	3774
Order est. (p)	1	1
Order bias (q)	2	2
BW est. (h)	8.741	8.741
BW bias (b)	8.741	8.741
rho (h/b)	1.000	1.000

Number of obs =	23132
BW type =	Manual
Kernel =	Triangular
VCE method =	NN

Outcome: Y. Running variable: X.

	Point Estimate	Robust Inference z-stat	P>|z|	[95% Conf. Interval]	
RD Effect	.27519	8.8880	0.000	.234469	.367133

Covariate-adjusted estimates. Additional covariates included: 1

Figure 5: rdlasso - Reduced form

For the outcome, the coefficient value is equal to 0.62631 percentage points with respect to the estimate of the point, the p-value is equal to zero, indicating high statistical significance, and the confidence interval corrected for robust bias 95% does not include zero.
As for the estimation of the treatment variable T, also for the outcome variable Y, the substantial difference between the two commands used `rdlasso` and `rdrobust` concerns the selection of the covariates: the first automatically selects only one, *icfes_ age_ 4*, corresponding to corresponding to the fourth power of the age variable, while the second includes five covariates (see Table A.1 in the Appendix).

The estimated fuzzy RD parameter, Figure 6, is equal to 0.4394, significant at 10%. Compared to the estimate obtained with `rdrobust`, the estimate of `rdlasso` remains substantially unchanged, as well as the confidence interval. This confirms that, through the automatic selection implemented by `rdlasso`, the substantive conclusions about the impact of the Ser Pilo Paga programme remain robust even in the context of high-dimensional covariates.

```
-------------------------------------------------------------
                   FUZZY ESTIMATOR
-------------------------------------------------------------
  RD Effect:                          0.4394
  Robust standard error:              0.0554
  z-statistic:                        7.9267
  p-value:                            0.0000
  95% Confidence Interval:  [   0.3307 ,    0.5480]
-------------------------------------------------------------
```

Figure 6: rdlasso - Covariate-adjusted estimates for fuzzy case

5 Conclusions

Covariates inclusion in RDD experiments is a widespread practice, to improve the precision of the estimates, following Calonico et al. (2019). Kreiss and Rothe (2023) extend this approach to high-dimensional settings under approximate sparsity by employing localized Lasso procedures for covariate selection.

The `rdlasso` command provides the Stata users with an easy way to implement ths methodology, combining the theoretical framework of Kreiss and Rothe (2023) with the computational efficiency of Python's Scikit-Learn library and Stata's `rdrobust` estimation environment. By allowing users to handle both sharp and fuzzy RDD within a single, unified interface, the command offers a flexible and reproducible solution for high-dimensional causal inference in Stata.

Future extensions may further integrate `rdlasso` with other ML frameworks developed in the spirit of double/debiased ML (Chernozhukov et al., 2018), such as the approach proposed by Noack et al. (2025), which generalizes localized regularization methods beyond Lasso. Such developments would not only expand the theoretical scope of the estimator, but also enhance the potential of Stata–Python integration for modern, high-dimensional econometric applications.

Appendix

A Sharp RDD and Fuzzy RDD with Low-Dimensional covariates

The table shows the results for the `rdrobust` command for the sharp RDD and fuzzy RDD application with and without covariates. The first two cases in the table concern the sharp RDD application, the resents the fuzzy RDD application

Table A.1: Sharp RDD and Fuzzy RDD with Low-Dimensional covariates

		Robust Inferenze			
	Point Estimate	z-stat	P> $\mid z \mid$	[95% Conf. Interval]	
`rdrobust` without covariates	3.0195	1.7759	0.076	-.309289	6.27577
`rdrobust` with covariates	3.108	2.0879	0.037	.193817	6.13153
`rdrobust` T X without covariates	.62483	43.1143	0.000	.595197	.65189
`rdrobust` Y X without covariates	.26904	10.0476	0.000	.220601	.327522
`rdrobust` T X with covariates	.62517	43.7480	0.000	.596446	.652396
`rdrobust` Y X with covariates	.27083	10.0485	0.000	.221952	.329516
`rdrobust` fuzzy without covariates 1	.61921	29.8862	0.000	.585361	.653051
`rdrobust` fuzzy without covariates 2	.43449	11.0262	0.000	.365964	.524194
`rdrobust` fuzzy with covariates 1	.6209	29.7333	0.000	.586823	.654971
`rdrobust` fuzzy with covariates 2	.43618	10.9901	0.000	.366835	.526076

The first two results come from the sharp RDD application without and with covariates. Looking at the coefficient of the treatment indicator, we can see that the percentage of women aged 15 to 20 who have completed high school increases by approximately 3,0195 percentage points with an Islamic victory. Looking at the robust bias-corrected confidence interval, we can see that it includes zero. The second result in the table shows how the effect of the estimated RD is quite similar to the unadjusted estimator. The most significant aspect is the reduction of the robust p-value, which ranges from 0.076 to 0.037.

The results in the third row of the table indicate that for the treatment variable T the point estimate is 0.62483 percentage points and the p-value is zero, indicating that the effect is statistically significant. Instead, the result in the fourth row reports the estimated value for the outcome variable Y of 0.26904 percentage points. Again, the p-value is zero, indicating that the effect is statistically significant. Applying the variables to the treatment value T the point estimate is 0.62517 percentage points with the p-value being zero, confirming the presence of a statistically significant effect. Doing the same for the outcome variable Y, the point estimate is 0.27083 percentage points and the p-value is zero, confirming, in this case too, the presence of a statistically significant effect.

The results in the last lines, however, come from using the fuzzy() option to calculate the fuzzy parameter. For the first stage results, the point estimate is 0.61921 with a p-value equal to zero, confirming the presence of a statistically significant effect. The results for Y, where the estimated treatment effect (LATE) is 0.43449, the p-value is zero and the confidence interval includes zero, also confirm the presence of a statistically significant effect.

B Installation guide

The following sections outline the essential steps for installing Python and configuring Stata to run the version of rdlasso released with this article. Further details are available through Chuck Huber on the Stata Blog.

B.1 Setting Stata to use Python

Stata allows for the use of Python by following a few simple steps. Typing *"python search"* in the stata command

```
. python search
```

- **If you have already installed Python on your PC:** the message *"Python environments found"* will be displayed with the path of the file where it is located;

```
. python search

Python environments found:
```

- **If Python has not been installed yet on your PC:** the following message will appear *"no Python installation found; minimum version required is 2.7. r(111)"*.

```
. python search

no Python installation found; minimum version required is 2.7.
r(111)
```

If the message in point two appears, you need to download Python from Python download website and install it.[4] Once the installation process has been completed, it will be possible to go back to Stata and run the *"python search"* command again to verify that the installation was successful.
If both Python and Anaconda are installed on your PC, you must choose which one to use by specifying the path with the *"set python_ exec"* command.

If you already have multiple versions of Python installed, to connect Stata to a specific version (3.8 or later), use the following command:

```
python set exec <pyexecutable> , permanently
```

where `pyexecutable` should be set depending on your operating system and where you installed Python.

It is possible to type *"python query"* to see which Python installation you are currently using.

B.2 Installation of Python packages

In the following we show how to install some of the required packages. An exactly similar procedure can be followed for any other package required by the Stata command and listed in the Section 3.1.

- **NumPy:** to check if the package is already installed just type the command "python which numpy" in Stata, the possible answers will be:
  - **Answer 1:** *"<module 'numpy' from ...>";*

    ```
    . python which numpy

    <module 'numpy' from ...>
    ```

  - **Answer 2:** *"Python module numpy not found r(601)"*

    ```
    . python which numpy

    Python module numpy not found
    r(601);
    ```

  In case of Answer 2 you need to proceed with the installation of the package by typing the command *"shell"* on Stata to open the Windows command prompt[5] and type *"pip -V"* to view the version and location of the pip programme. Then, use the command *"pip install*

[4] We recommend using the 64-bit installer. We also remind the user that it is possible to download the appropriate installer from Anaconda website.

[5] The same command will also open a terminal in Mac or Linux operating systems.

numpy" in the command prompt to download and install the NumPy package.
To verify the correct installation, just type again the command *"python which numpy"* in Stata, in this case the Answer 1 will be shown.

- **Pandas:** to install this package you must simply follow the same procedure as for the case of NumPy typing *"pip install pandas"* in the prompt.
- **Scikit-Learn:** to install this package you must simply follow the same procedure as for the case of NumPy typing *"pip install sklearn"* in the prompt.

Then, use the command *"pip install"* to verify that numpy, pandas, scikit-learn are installed.

Bibliography

Armstrong, T. B. and Kolesár, M. (2018). A simple adjustment for bandwidth snooping. *The Review of Economic Studies*, 85(2):732–765.

Belloni, A. and Chernozhukov, V. (2013). Least squares after model selection in high-dimensional sparse models.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2):608–650.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, 17(2):372–404.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3):442–451.

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

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2019). *A practical introduction to regression discontinuity designs: Foundations*. Cambridge University Press.

Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2024). *A Practical Introduction to Regression Discontinuity Designs: Extensions*. Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge University Press.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., and Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters.

Cox, C. (1990). Fieller's theorem, the likelihood and the delta method. *Biometrics*, pages 709–718.

Imbens, G. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of economic studies*, 79(3):933–959.

James, G., Witten, D., Hastie, T., and Tibshirani, R. (2013). *An introduction to statistical learning: with applications in R*, volume 103. Springer.

Kreiss, A. and Rothe, C. (2023). Inference in regression discontinuity designs with high-dimensional covariates. *The Econometrics Journal*, 26(2):105–123.

Londoño-Vélez, J., Rodríguez, C., and Sánchez, F. (2020). Upstream and downstream impacts of college merit-based financial aid for low-income students: Ser pilo paga in colombia. *American Economic Journal: Economic Policy*, 12(2):193–227.

Meyersson, E. (2014). Islamic rule and the empowerment of the poor and pious. *Econometrica*, 82(1):229–269.

Noack, C., Olma, T., and Rothe, C. (2025). Flexible covariate adjustments in regression discontinuity designs. *arXiv preprint arXiv:2107.07942*.

Shabbir, J. (2021). An introduction to statistical learning with applications in r. *Statistical Theory and Related Fields*.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288.

Chapter three

Grandparental Retirement and the Fertility Decisions of Adult Children: Evidence from Europe

Marianna Nitti[+] Marco Ventura[+]

[+] *Sapienza, University of Rome*

Abstract

This paper estimates the causal effect of grandparental retirement on the fertility of their adult children in a pan-European context. To test the generalizability of this relationship across nations with different welfare systems, we employ a fuzzy Regression Discontinuity Design (RDD). Using panel data from the Survey of Health, Ageing, and Retirement in Europe (SHARE), our identification strategy exploits the exogenous and discontinuous variation in the probability of retirement at statutory pension eligibility ages. Our findings confirm a strong and statistically significant first stage across all countries, indicating that eligibility substantially increases the likelihood of retirement. However, in contrast to previous single-country studies, we do not find evidence of a positive immediate fertility response following parental retirement. The short-term effects at $t+1$ are generally null or negative, while medium-term estimates at $t+2$ and $t+3$ reveal persistent negative fertility responses, particularly in Mediterranean and Eastern European countries. The results suggest that the negative income effect associated with retirement dominates the positive time effect generated by increased grandparental childcare availability. Additional evidence from pre-treatment periods highlights substantial forward-looking behaviour, indicating that adult children strategically adjust fertility timing in anticipation of parental retirement. Overall, the findings reveal that the demographic consequences of retirement critically depend on welfare regimes, intergenerational financial dependence, and the structure of family support systems.

Keywords: Fertility, Gender, Retirement, Fuzzy Regression Discontinuity Design

1 Introduction and relevant literature

Europe is experiencing a profound demographic transition, characterised by persistently low birth rates and a progressively ageing population. According to Eurostat, the European fertility rate reached a record low of 1.38 live births per woman in 2023, well below the replacement level of 2.1. This demographic pressure is exacerbated by economic uncertainty and instability in the labour market, which heavily influence the fertility decisions of younger generations (Cavallini, 2024). Furthermore, these national dynamics are compounded by significant regional disparities, with fertility rates varying greatly based on local socioeconomic contexts (Salvati et al., 2020).

A key driver of this trend is the timing of fertility decisions. The mother's age at first birth has become a critical variable in determining final family size, highlighting the tension between the biological clock and modern life-course trajectories characterised by prolonged education and delayed labour market entry. Delaying childbearing not only narrows the reproductive window, but also reduces the likelihood of conception, a factor identified by health institutions as the leading cause of infertility (Zanardo et al., 2025). This trend is well documented: In all OECD countries, including Italy, the average age at first birth rose from 26.5 years in 2000 to 29.5 years in 2022 (Eurostat, 2025).

This demographic shift, resulting in a pronounced imbalance in the population's age structure, places significant strain on the sustainability of welfare systems and challenges the intergenerational balance.

In this context, intergenerational transfers within the family become critically important. The older population, in particular, often represents a pillar of informal welfare by dedicating considerable time to supporting their children, for example, through grandchild care. This observation has led researchers to investigate the link between grandparental time availability and their children's fertility. However, establishing a causal link is challenging due to endogeneity concerns. To overcome this, recent literature has turned to quasi-experimental settings, hypothesizing that pension system reforms provide an exogenous shock to the availability of informal childcare.

A first stream of research has analysed the effects of reforms that postpone retirement age, showing how they can negatively impact fertility. The study by Ilciukas (2023) on a Dutch reform, for example, highlights how forced postponement of retirement reduces the likelihood that adult children have a child, especially for families living in close geographical proximity. Similarly, Akyol and Atalay (2025) show that longer working lives for grandmothers reduce the time they can dedicate to caring for their grandchildren, which in turn negatively impacts fertility, an effect that is particularly pronounced in families with limited economic resources.

In parallel, the literature has explored the reverse causality channel by examining the impact of a grandchild's birth on grandparents' work decisions, confirming the strength of the endogenous link. Studies such as those of Rupert and Zanella (2018) and Backhaus and Barslund (2021) demonstrate that becoming a grandmother significantly reduces the probability of working for older women. This effect is negligible for men and is stronger in countries with limited public childcare provision.

Complementing these empirical studies, theoretical frameworks such as the overlapping generations (OLG) model by Baldi et al. (2025) formalise these dynamics. In this model, grandparental care is defined as a "time transfer" that allows adult children to invest greater resources in their children's education, generating potential benefits for the entire community.

In this context, characterised by complex endogenous interactions, studies based on a regression discontinuity design (RDD) offer the most robust methodological solution for isolating the causal effect. The crucial study in this vein is that of Eibich and Siedler (2020), which analyse the case of Germany. Using retirement age thresholds as an exogenous discontinuity, the authors show that when parents reach retirement eligibility, the probability of conception for their adult children increases significantly, providing direct and causal evidence of the link between retirement and fertility.

Building directly on this latter work, our study extends the analysis of Eibich and Siedler (2020) to a pan-European context. We adopt the same conceptual framework by analysing a three-generation family structure composed of grandparents (first generation), adult children (second generation), and potential grandchildren (third generation). To this end, we employ an RDD using data from the Survey of Health, Ageing, and Retirement in Europe (SHARE), leveraging its panel component to link data on the first generation with their adult children. Our identification strategy exploits the discontinuous increase in the probability of retirement that occurs when an eligibility threshold is reached. This threshold is defined specifically for each country and year, allowing us to isolate the effect of grandparents' retirement on their children's fertility. The objective is to test whether the positive causal effect identified in Germany holds in countries with different social systems and welfare systems.

Our RDD estimators reveal a robust and positive first stage, confirming that reaching pension eligibility markedly increases the propensity to retire across all European welfare regimes. However, when focusing on the immediate short-term horizon ($t + 1$), the second-stage analysis reveals that this surge in grandparental time availability does not translate into a prompt increase in the fertility of adult children. In contrast to previous single-country studies, our pan-European estimates show generally null or slightly negative short-term effects.

Exploring medium-term dynamics further strengthens this result: the negative fertility effects persist at $t + 2$ and $t + 3$, particularly in Mediterranean and Eastern European countries. Additional evidence from pre-treatment periods also reveals substantial forward-looking behaviour, suggesting that adult children strategically adjust fertility timing in anticipation of parental retirement. Overall, our findings suggest that the positive "time effect" generated by increased grandparental childcare availability is largely offset — and in some contexts dominated — by the negative "income effect" associated with retirement and the reduction in intergenerational financial transfers.

The remainder of the paper is structured as follows. Section 2 presents the dataset, and Section 3 describes the identification strategy. The results are discussed in Section 4. Section 5 concludes and provides some policy recommendations.

2 Data

2.1 Description of the dataset and sample construction

This study draws on data from the Survey of Health, Ageing, and Retirement in Europe (SHARE), specifically waves 2 (SHARE-ERIC (2024a)), 4 (SHARE-ERIC (2024b)), 5 (SHARE-ERIC (2024c)), 6 (SHARE-ERIC (2024d)), 7 (SHARE-ERIC (2024e)), 8 (SHARE-ERIC (2024f)), and 9 (SHARE-ERIC (2024g)). SHARE provides rich panel data on individuals aged 50 and older in numerous European countries, covering demographics, health, employment, pensions, and social networks. Our analysis includes nineteen countries from waves 2 through 9. We specifically selected countries where the institutional pension landscape allows for the identification of a clear discontinuity in retirement probability, ensuring a robust first stage for the RDD estimation. Crucially for our analysis of heterogeneous effects, we classify these countries into distinct welfare regimes (Continental, Mediterranean, Nordic, and Eastern) following established sociological and economic classifications (e.g., Stella (2017)). To determine retirement eligibility, we reconstruct the specific pension rules for each country and year using information from the Mutual Information System on Social Protection (MISSOC), as detailed in the Appendix **??**.

The analysis is conducted on a child-year panel constructed from SHARE intergenerational linkages. Each adult child-year observation can be linked separately to the retirement status and pension eligibility conditions of the biological father and biological mother, allowing the estimation of gender-specific grandparental retirement effects. Our main sample consists of adult children for whom data on both parents are available. Additionally, we conduct the analysis on two sub-samples, focusing exclusively on the retirement of the mother and the father, respectively. We excluded several groups from the sample: individuals who did not report a retirement year; those who retired for reasons other than meeting statutory pension requirements; and elderly individuals without children. More specifically, to isolate retirement transitions more plausibly driven by statutory pension eligibility, we exclude individuals reporting retirement due to health shocks, family reasons, firm downsizing, or coordinated retirement decisions. Finally, to focus on the relevant life stages, we restrict the sample to elderly parents aged 40-80 and adult children aged 20-50, aligning with the primary fertility window.

Following these restrictions, the final sample comprises $91,553$ adult children and covers the period from 2004 to 2022.

2.2 Retirement, pension eligibility, and age

Our treatment variable, retirement status, is constructed from the self-reported employment status of the elderly parents. We treat retirement as an absorbing state, meaning that individuals are considered to remain retired for all subsequent periods after their initial transition.

An individual's age is calculated precisely in months using their interview and birth dates. Pension eligibility is determined for each individual based on their gender and the specific rules in effect in a given year, which may also depend on their birth cohort.

For each person-year observation, we compute the eligibility age for both early and old-age retirement. In general, we define the effective eligibility threshold as the minimum of these two ages, reflecting that individuals can retire as soon as they meet the earliest possible legal criterion. However, to preserve the validity of our Regression Discontinuity Design, for countries where early retirement is strictly conditional on stringent and unobservable contribution histories (e.g., 30 or 40 years of continuous contributions), we exclude the early retirement pathway. For these specific cases, the effective eligibility threshold is anchored exclusively to the statutory old-age retirement age, ensuring a sharp and exogenous discontinuity in the retirement probability.

The running variable (X), which measures the distance to the eligibility threshold, is then calculated as the difference between the individual's current age and their effective eligibility age, both measured in months. The running variable is measured in months, substantially increasing support points around the cutoff. This variable is normalised such that it takes negative values before an individual reaches the threshold, zero at the threshold, and positive values thereafter. Finally, we construct our instrumental variable (Z), an indicator of eligibility status. The instrument is equal to 1 if the running variable is zero or positive ($X \geq 0$) and 0 otherwise.

2.3 Outcome variables and descriptive statistics

The primary focus of our analysis is the fertility of adult children. SHARE does not provide complete annual fertility histories for adult children. Therefore, childbirth events are reconstructed longitudinally using updates in the reported birth year of the youngest grandchild. Our main outcome variable is a binary indicator capturing whether a child is born in the year following the parent's retirement ($t + 1$). To capture the more complex and delayed timing of fertility decisions, we also construct indicators for births occurring two ($t + 2$) and three years ($t + 3$) after retirement. However, to maintain a sharp focus on the immediate demographic response in the main text, the medium-term results are extensively documented and discussed in the Appendix.

This forward-looking approach to $t + 1$ is crucial as it accounts for the nine-month biological gestation period, allowing us to evaluate the immediate reaction to the retirement shock. Focusing on this short-term horizon enables us to test whether the sudden increase in grandparental time availability translates into a prompt increase in fertility, or if fertility decisions are instead rigid in the very short term, possibly due to the offsetting negative income effect associated with retirement.

Furthermore, to validate our causal identification strategy, we formally test for potential anticipation effects by examining fertility changes in the years preceding the parent's retirement, specifically at $t-1$, $t-2$, and $t-3$. This serves as a rigorous falsification (placebo) test: if fertility decisions are primarily driven by the actual retirement event rather than pre-existing trends or forward-looking anticipation, we expect to find no significant effects in the pre-treatment periods.

Table 1 presents the summary statistics for the sample of adult children, disaggregated by gender (sons and daughters) and for the full sample. Panel A details the outcome variables, focusing primarily on fertility events, while panel B reports the descriptive statistics for a set of key covariates.

Table 1: Descriptive statistics: Outcome variables and covariates for adult children

	Variable	**Sons**	**Daughters**	**Adult children - all**
A	birth	0.019	0.022	0.020
	(std. dev.)	(0.137)	(0.145)	(0.141)
	birth (t+1)	0.018	0.021	0.019
	(std. dev.)	(0.134)	(0.142)	(0.138)
	birth (t+2)	0.018	0.020	0.019
	(std. dev.)	(0.132)	(0.141)	(0.137)
	birth (t+3)	0.018	0.020	0.019
	(std. dev.)	(0.132)	(0.139)	(0.136)
B	n of children	0.862	1.028	0.942
	(std. dev.)	(1.125)	(1.168)	(1.149)
	gender	1	0	0.513
	(std. dev.)	(0)	(0)	(0.500)
	age	34.710	34.600	34.656
	(std. dev.)	(8.086)	(8.122)	(8.103)
	distance from parents	0.282	0.242	0.262
	(std. dev.)	(0.450)	(0.428)	(0.440)
	marital status	0.308	0.338	0.323
	(std. dev.)	(0.462)	(0.473)	(0.468)
	partner	0.246	0.295	0.268
	(std. dev.)	(0.431)	(0.456)	(0.443)
	contacts with parents	0.691	0.727	0.708
	(std. dev.)	(0.462)	(0.446)	(0.455)
	employment status	0.804	0.717	0.762
	(std. dev.)	(0.397)	(0.450)	(0.426)
	education	0.752	0.793	0.772
	(std. dev.)	(0.432)	(0.405)	(0.420)

Note: The covariates presented in Panel B are constructed as binary indicators. Specifically: *distance from parents* is equal to 1 if the adult child lives in close proximity to the parents (i.e., within 25 km); *marital status* equals 1 if the child is married or in a registered partnership; *partner* equals 1 if the child currently has a partner; *contacts with parents* is equal to 1 if the child is in contact with the parents at least once a week (including daily and several times a week); *employment status* equals 1 if the child is currently employed (full-time, part-time, or self-employed); and *education* is equal to 1 if the child has attained at least a high school diploma or equivalent. Number of observation (child-years): 1.153.527. Number of unique adult children: 87.538.

Some interesting insights emerge from Panel A. As expected, the probability of giving birth in any given year is relatively low, approximately 2.0% for the whole sample (*birth*). This probability remains remarkably stable when considering individual births in the following year ($t+1$, 1.9%), as well as two ($t+2$, 1.9%) and three years ($t+3$, 1.9%) into the future. A notable gender-based pattern can be observed: daughters consistently show a slightly higher probability of giving birth across all immediate and future time horizons compared to sons. Consequently, the daughters in the sample have, on average, a higher number of children (1.03) than the sons (0.86).

Panel B provides an overview of the characteristics of the adult children in the sample. The average age is nearly identical for sons (34.7) and daughters (34.6), with a total sample mean of approximately 34.7 years. Regarding education, daughters on average report a higher level of educational attainment than sons (79.3% vs 75.2%). In terms of family ties, an interesting distinction appears: daughters report a higher frequency of contact with their parents (72.7%) compared to sons (69.1%). Differences in marital, partnership, and employment statuses are also observable between the two groups, with sons showing a significantly higher employment rate (80.4%) than daughters (71.7%).

In general, these descriptive statistics provide a basic overview of the sample, highlighting subtle but potentially significant gender differences in fertility patterns, education, and family dynamics. These characteristics, particularly the gendered nature of parental contacts and employment, will be crucial to consider in the subsequent causal analysis.

Table 2: Descriptive statistics: Covariates for elderly parents

Variable	**Grandfathers**	**Grandmothers**
marital status	0.654	0.540
(std. dev.)	(0.476)	(0.498)
born in country of interview	0.901	0.910
(std. dev.)	(0.298)	(0.286)
citizen of the country of interview	0.956	0.967
(std. dev.)	(0.204)	(0.177)
highest degree	0.348	0.341
(std. dev.)	(0.476)	(0.474)
years of education	11.531	11.531
(std. dev.)	(4.441)	(4.441)
n of children	2.858	2.844
(std. dev.)	(1.371)	(1.400)
cohabiting	0.526	0.423
(std. dev.)	(0.499)	(0.494)

Note: Number of unique grandfathers: 19660. Number of unique grandmothers: 23365.

Table 2 presents the summary statistics for the covariates of the elderly parents in the sample, disaggregated by gender into grandfathers and grandmothers.

The data reveal several notable differences between the two groups. In terms of education, both groups show remarkably similar levels of attainment, with an average of about 11.5 years of education, although grandfathers have a marginally higher likelihood of holding a higher degree (34.8% vs. 34.1%).

Interesting patterns also emerge in the household and family structure. Grandfathers are more likely to cohabit than grandmothers (52.6% vs. 42.3%), a finding that is consistent with the observed differences in marital status and may reflect broader demographic trends related to life expectancy and remarriage rates for men. As expected, the average number of children is almost identical for both groups (approximately 2.8). Finally, variables related to origin, such as citizenship and country of birth, show minimal differences, indicating a homogeneous sample in this regard.

3 Identification strategy

3.1 The model

We estimate the effect of grandparents' retirement on their adult children's fertility by exploiting the discontinuity in the probability of retirement induced by statutory pension eligibility. Our identification strategy builds on the exogenous variation in retirement incentives generated by age-based eligibility rules, which define precise cutoffs for the earliest and statutory pension ages. In most European countries, these thresholds are determined by national legislation; for instance, the statutory retirement age is fixed by law, while early retirement schemes allow individuals to retire several years earlier with reduced benefits.

These institutional rules create discontinuities in retirement eligibility that are exogenous to individual fertility decisions. Around these thresholds, retirement behaviour responds discretely to changes in eligibility, while other determinants of intergenerational fertility decisions are expected to vary smoothly.

The sample comprises interviewed individuals aged 40 years and older, covering the years 2004 to 2022. Let us denote by X_{it} the running variable (normalised to zero) measuring the distance from/to the eligibility of the elderly parent i in year t. Consequently, let $Z_{it} = I(X_{it} \geq 0)$ be the corresponding eligibility status of the elderly parent i in year t; whenever $Z_{it} \geq 0$, the elderly parent i is eligible for retirement in year t. The treatment variable, R_{it}, is equal to 1 for retired elderly parents and 0 otherwise. The outcome of interest, $Y_{i,t+1}$, measures the fertility status of the adult child i (that is, the probability that a birth occurs in the year following the interview, $t + 1$).

At the cutoff (e.g., $X_{it} = 0$) we will observe a discontinuous change in the probability of being treated. The probability of retirement does not jump from zero to one because retirement is not compulsory; it is an option in the hands of the worker, and some individuals may decide to continue working even if they have met their minimum pension requirements. Furthermore, eligibility rules are subject to some limited but not excludable exceptions. In particular, disabled people, war veterans, recipients of reversibility pensions, and other peculiar cases are characterised by shorter working periods compared to the common minimum requirement. In other words, it could happen that some individuals retire early compared to the standard eligibility rule, due to legal exceptions. We are not able to fully control all the factors that alter the general requirements as established by law. It follows that the setting just presented is a well-known fuzzy RDD in which the discontinuity in the probability of retirement helps to overcome the problems arising from endogenous selection in retirement.

Specifically, our setting is a fuzzy case with double non-compliance, that is, on the one side we have ineligible retirees (e.g. disabled people, veterans, etc.) who are actually retired, and on the other side we have eligible individuals who are still working. In this setup RD can recover the local average treatment effect (LATE), the effect on compliers at the cutoff, where compliers are defined as individuals whose treatment status changes as we move the value of X_{it} from just

to the left of the cutoff to just to the right of it (Angrist and Pischke, 2009). This is like saying that compliers are individuals who receive treatment if their score is above the threshold but do not receive it if their score is below the threshold.
A valid fuzzy RD design relies on three main assumptions. The first assumption requires that conditional on R_{it}= 0,1 there is *no discontinuity* in the regression functions at the cutoff. This implies that in the absence of retirement, the expected fertility of adult children would evolve smoothly at the cutoff point. In other words, "all other factors" driving fertility must be continuous at the threshold (see, e.g. Hahn et al., 2001).
The second assumption requires that the potential treatment $R_{it}(X)$ must be *monotonic* with respect to the score X at the threshold c. This implies that there are no "defiers," - that is, individuals who would have received the treatment if their score were below the threshold but would not have received it if their score were above the threshold; and there is at least one complier.
The third assumption, not always explicitly stated, is that there is a discontinuity in the probability of receiving treatment at the cutoff, commonly referred to as the existence of a *first stage*. This means that the treatment probability must show a jump at the cutoff c.
The panel dimension allows us to control for time invariant unobserved individual heterogeneity through fixed effects, but does not alter the core identifying assumptions of the fuzzy RDD.

From a strictly empirical point of view, RDD is implemented using a two-stage least squares (2SLS) estimation where the instrument is eligibility Z_{it}. The estimation includes individual fixed effects and year fixed effects to account for unobserved heterogeneity. The first stage equation is as follows:

$$R_{it} = \alpha_1 + \theta Z_{it} + \beta_1 f\left(X_{it}\right) + \gamma_i + \kappa_t + \nu_{it} \tag{3.1}$$

where $f(X_{it})$ is a power series of X_{it}, θ captures the jump in the probability of retirement induced by achieving eligibility, i.e. the proportion of compliers in the sample, γ_i and κ_t are individual and year fixed effects, and ν_{it} is a random error term. From this stage, we obtain the predicted values for retirement, $\hat{R}_{it}$.

The second stage equation then uses these predicted values to estimate the causal effect on the outcome:

$$Y_{i,t+1} = \alpha_2 + \lambda \hat{R}_{it} + \beta_2 f\left(X_{it}\right) + \tau_i + \omega_t + \epsilon_{it} \tag{3.2}$$

where $\hat{R}_{it}$ is the predicted retirement status from the first stage, λ is the causal effect of retirement on the outcome of interest, τ_i and ω_t are individual and year fixed effects, and ϵ_{2it} is a random error term.

Other controls can be easily added to the equations (3.1) and (3.2). The estimation is carried out at the cutoff, meaning in a narrow bandwidth around the cutoff. Following Eibich and Siedler (2020), our baseline specification employs a 120-month (10-year) bandwidth. The amplitude of the bandwidth is an empirical question, and we will test the robustness of our choice.

Table 3: Further descriptives for grandfathers and grandmothers

Variable	Grandfathers	Grandmothers
age (in months)	765.685	742.106
	(95.010)	(98.589)
X	17.212	6.780
	(104.674)	(105.833)
Z	0.545	0.502
	(0.498)	(0.500)
R	0.501	0.450
	(0.500)	(0.498)

Table 3 presents the summary statistics for the key variables used in our RDD design, with data broken down for grandfathers and grandmothers. A clear pattern emerges from the data. The grandfathers in the sample are on average older than the grandmothers (approximately 766 vs. 742 months). This age differential is logically reflected in the other variables. The mean value of the running variable (X), which represents the distance from the pension eligibility threshold, is higher for men (approximately 17 vs. 7 months). This indicates that, on average, they are further past the point of eligibility. Consequently, the share of grandfathers who are eligible (Z, 54.5% vs. 50.2%) and the share who are actually retired (R, 50.1% vs. 45.0%) are higher than the corresponding figures for grandmothers. The running variable is measured in months, substantially increasing the number of support points around the eligibility threshold.

3.2 Testing the assumptions

A valid fuzzy RD design requires the continuity of the conditional regression of the expected outcome at the cutoff. This assumption implies that in the absence of treatment, the potential outcomes and all other relevant predetermined covariates should not jump at the threshold (Calonico et al., 2019).

In our framework, the assignment to treatment relies on exact birth dates and strictly exogenous institutional pension rules. Because individuals cannot retroactively alter their date of birth to manipulate their pension eligibility, predetermined time-invariant covariates (such as gender, country of birth, etc.) and covariates that evolve predictably over time (e.g., age) are expected to be continuous at the cutoff point. Therefore, there is no reason to believe that the units just above and just below the cutoff differ systematically in their predetermined characteristics.

An issue strictly related to *continuity* is the potential manipulation of the running variable. Indeed, if manipulation takes place, there will no longer be randomly assigned treated/controls at the cutoff, and the running variable will not be distributed continuously. As mentioned, the manipulation hypothesis is highly unlikely in our context because qualifying for retirement would require individuals to manipulate their exact age and birth records, which is virtually impossible.

To formally rule out this concern, we implement the density test proposed by McCrary (2008).

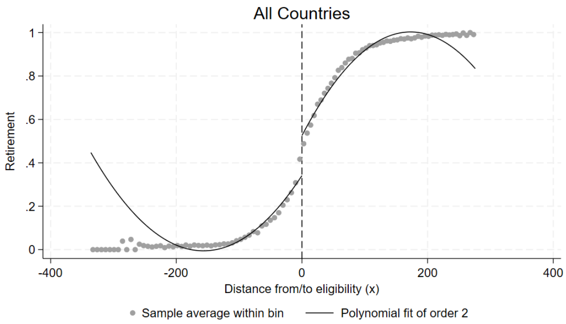
In accordance with our expectations, the test returns a P-value for the null hypothesis of no manipulation equal to 0.4418 for grandfathers and 0.3381 for grandmothers. Visual evidence for the continuity of the running variable's distribution, along with the McCrary (2008) test plots, are provided in Figures A.1 and A.2.

As far as the *first stage* assumption is concerned, visual evidence strongly supports the relevance of our instrument. Figure 1 graphically proves the existence of a robust first stage for the pooled sample of all countries, showing a sharp and positive jump in the probability of being treated (i.e., retired) exactly at the eligibility threshold ($X = 0$) for both grandfathers and grandmothers.

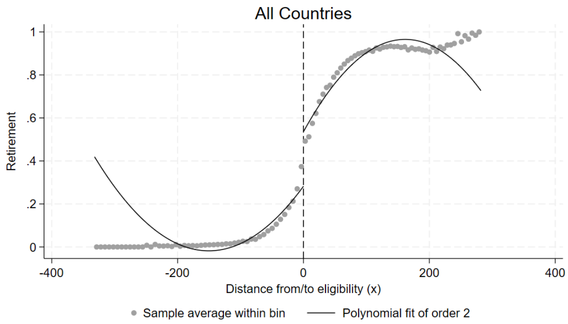
To explore the institutional heterogeneity that is central to our analysis, we further disaggregate the first-stage plots by welfare regime. Figures 1 and 2 illustrate the discontinuity for the Continental, Mediterranean, Nordic, and Eastern regimes. Across all four welfare systems, the probability of retirement increases discontinuously and markedly when elderly parents reach their specific pension eligibility age.

A closer inspection of these plots reveals two important patterns that validate our empirical strategy. First, the probability of retirement to the left of the cutoff is strictly positive (ranging approximately between 20% and 40%), reflecting early retirement exceptions, disability pensions, or alternative exit routes from the labor market. Second, the probability does not jump perfectly to one to the right of the cutoff, indicating that a share of eligible individuals chooses to continue working. This double non-compliance perfectly visually justifies the use of a fuzzy Regression Discontinuity Design.

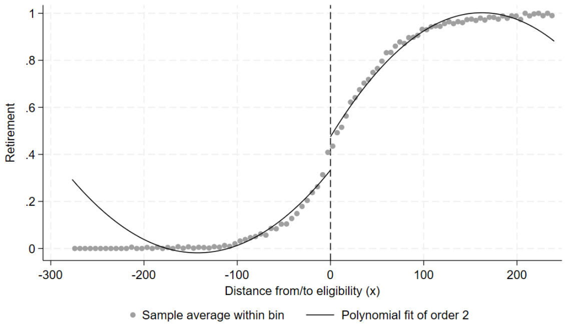
Figure 1: First stage plot - Grandfathers and Grandmothers



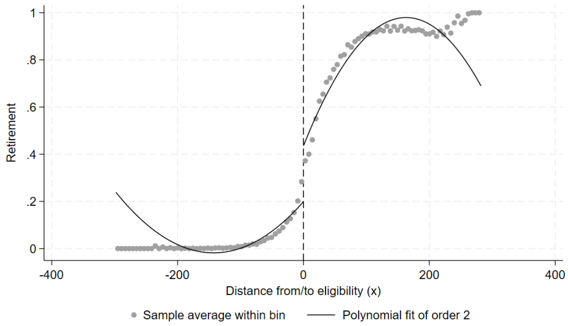
(**a**) First Stage - Grandfathers All countries



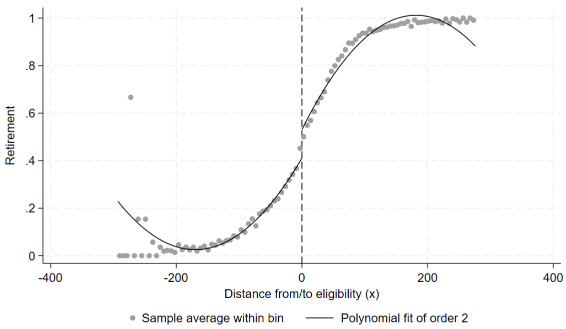
(**b**) First Stage - Grandmothers All countries



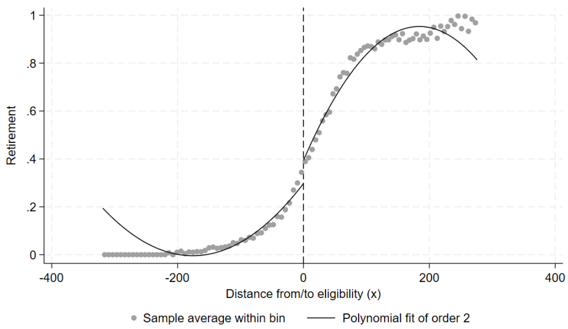
(**c**) First Stage - Grandfathers Continental



(**d**) First Stage - Grandmothers Continental

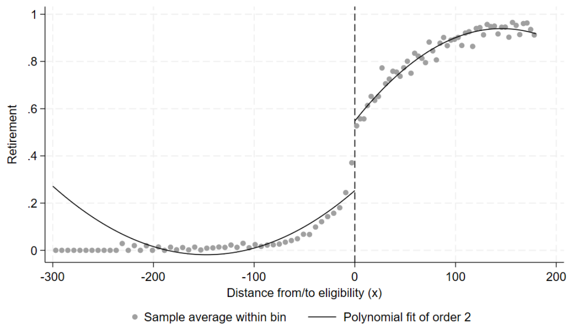


(**e**) First Stage - Grandfathers Mediterranean

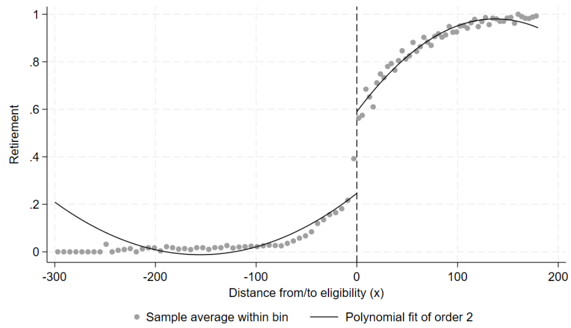


(**f**) First Stage - Grandmothers Mediterranean

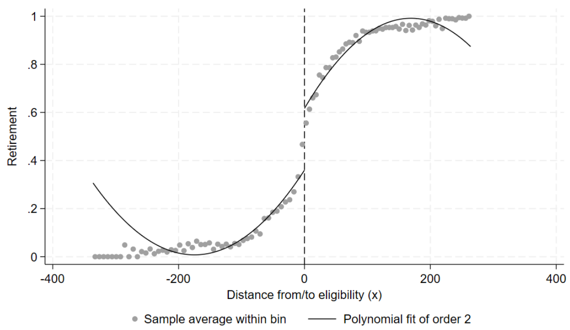
Figure 2: First stage plot - Grandfathers and Grandmothers



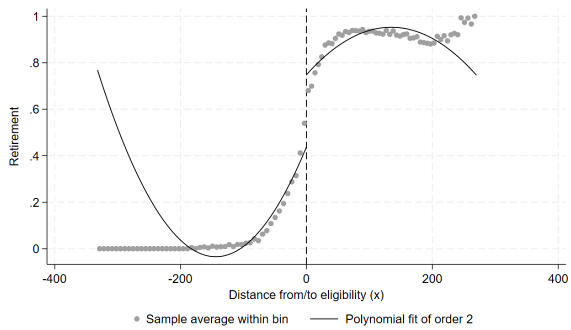
(**g**) First Stage - Grandfathers Nordic



(**h**) First Stage - Grandmothers Nordic



(**i**) First Stage - Grandfathers Eastern



(**l**) First Stage - Grandmothers Eastern

Note: The figure reports the jump in the probability of treatment at the cutoff. On the y-axis of each panel there is the probability of retirement, while the x-axis reports time from/to eligibility, centered at zero. The black line represents the underlying regression function, the dots represent the average probability at a given relative time, the dashed line is the cutoff. The bins are computed according to the evenly-spaced method as proposed by Calonico et al. (2015).

4 Results

4.1 First Stages

We move on now to the estimation of the parameters of the causal effect of Z_{it} on R_{it}; that is, the first stage that provides us with relevant information on the effect of eligibility on retirement at the cutoff, as well as the proportion of compliers and the take-up rate of the policy.

Following the approach in Eibich and Siedler (2020), for our main specification we employ a local quadratic polynomial for the age trend within a 120-month (10-year) bandwidth. This choice ensures that our results are directly comparable to this benchmark study. The use of a quadratic specification is further supported by the literature.[1]

Table 4 reports the results for our first-stage benchmark specification across the different European welfare regimes, disaggregated by the gender of the elderly parent and the adult child.

[1] Porter (2003) claims that local polynomials of lower order than quadratics suffer from boundary bias, while Gelman and Imbens (2019) show that local polynomials of higher order than quadratics tend to be inaccurate and overfit the data.

Table 4: First stage by groups of countries

Group of Countries		Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Eligibility	0.231***	0.231***	0.232***
	dummy (Z)	(0.003)	(0.004)	(0.004)
Continental	Eligibility	0.220***	0.281***	0.221***
	dummy (Z)	(0.004)	(0.006)	(0.006)
Mediterranean	Eligibility	0.142***	0.143***	0.141***
	dummy (Z)	(0.007)	(0.010)	(0.010)
Nordic	Eligibility	0.342***	0.334***	0.349***
	dummy (Z)	(0.009)	(0.012)	(0.012)
Eastern	Eligibility	0.284***	0.297***	0.271***
	dummy (Z)	(0.008)	(0.011)	(0.011)
B. Grandmothers				
All Countries	Eligibility	0.289***	0.285***	0.294***
	dummy (Z)	(0.003)	(0.004)	(0.004)
Continental	Eligibility	0.286***	0.280***	0.292***
	dummy (Z)	(0.004)	(0.006)	(0.006)
Mediterranean	Eligibility	0.120***	0.126***	0.114***
	dummy (Z)	(0.007)	(0.010)	(0.010)
Nordic	Eligibility	0.379***	0.380***	0.378***
	dummy (Z)	(0.009)	(0.012)	(0.012)
Eastern	Eligibility	0.341***	0.332***	0.350***
	dummy (Z)	(0.005)	(0.008)	(0.008)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for adult children and elderly parents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***\ p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$.

The first-stage results reveal a large, positive, and highly statistically significant jump in the probability of retirement when parents cross the eligibility threshold across all specifications. For the pooled sample of all countries, reaching pension eligibility increases the probability of retirement by 23.1 percentage points for grandfathers and 28.9 percentage points for grandmothers. The extremely low standard errors confirm the strong relevance of the instrument and rule out concerns related to weak instruments.

From a Local Average Treatment Effect (LATE) perspective, this coefficient represents the exact share of compliers at the cutoff. The observed variation thus suggests that the take-up rate of retirement policies at the earliest point of eligibility differs significantly across European welfare states.

Indeed, when disaggregating the analysis by welfare regime, an interesting pattern of institutional heterogeneity emerges. The Nordic regime displays the highest proportion of compliers, with a massive jump of 34.2 percentage points for men and 37.9 percentage points for women. The Eastern regime shows similarly robust effects, with a jump of 28.4 percentage points for grandfathers and 34.1 percentage points for grandmothers. The Continental regime exhibits a slightly lower compliance rate, hovering around 22.0% for men and 28.6% for women. Conversely, the Mediterranean regime displays a markedly lower first-stage jump, equal to 14.2 percentage points for men and just 12.0 percentage points for women. This more modest compliance rate in Southern Europe is entirely consistent with the structural characteristics of these labor markets, which historically featured highly fragmented working careers and widespread use of alternative early retirement or disability pathways prior to the statutory age.

Nevertheless, the discontinuity remains highly significant ($p < 0.01$) across all regimes, validating our identification strategy and demonstrating that age-based eligibility acts as a powerful shock to labor supply across the entire European context.

4.2 Fuzzy RD parameters

Again, we first present a visual analysis of the effect under scrutiny, and then move on to a more formal approach. Figure 3 and 4 provide a graphical representation of the Regression Discontinuity design for our main outcome, births at $t + 1$, for the pooled sample, while Figures 5 - 12 disaggregate the visual evidence by European welfare regime. In each plot, the points represent local averages of births binned over the running variable. The solid vertical line marks the eligibility cutoff, where we can visually inspect for a discontinuity in the outcome. For completeness, the detailed visual analysis for individual countries at $t + 1$, as well as the corresponding RD plots for births at $t + 2$ and $t + 3$, are provided in the Appendix.

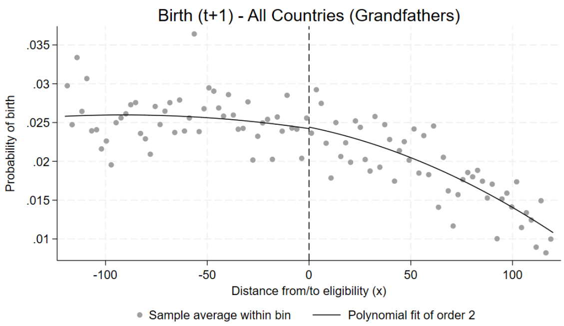


Figure 3: Birth (t+1) - Sons - All countries

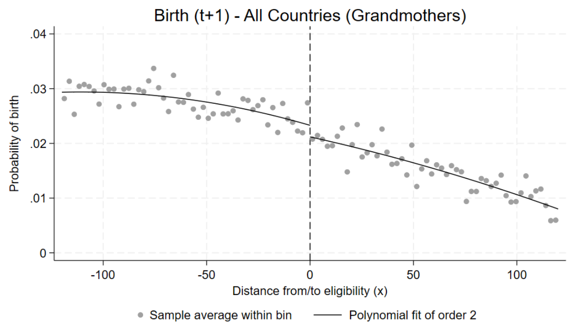


Figure 4: Birth (t+1) - Daughters - All countries

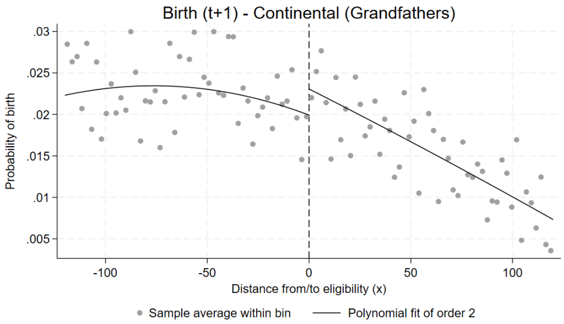


Figure 5: Birth (t+1) - Sons - Continental

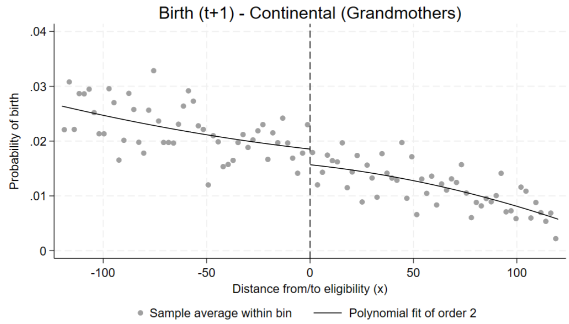


Figure 6: Birth (t+1) - Daughters - Continental

A visual inspection of the RD plots across the different welfare regimes offers preliminary insights into the fertility response to parental retirement. In contrast to the massive discontinuities observed in the first stage, the visual jumps in the probability of a birth at $t + 1$ are markedly modest. While some regimes reveal a slight positive jump at the cutoff, suggesting a potential immediate increase in the probability of a birth, the visual evidence is not uniform, with other regimes showing a visually null or even noisy discontinuity.

This heterogeneity is an interesting finding in itself and can be plausibly explained by the rigidities of short-term fertility adjustments. The absence of a sharp positive jump at $t + 1$ should therefore not be interpreted as a lack of effect, but rather as perfectly consistent with our theoretical framework: the positive effect driven by the sudden availability of grandparental time is temporarily muted by the biological gestation lag and offset by the immediate negative income shock associated with retirement. Thus, the fertility response is delayed beyond the first year, materializing more clearly in subsequent periods (as documented in the Appendix for $t + 2$ and $t + 3$).

Table 5 presents the main Reduced Form and Second Stage estimates of the causal effect of parental retirement on the probability that adult children will have a child in the immediately following year (t+1). The econometric analysis formally confirms the visual intuition: rather than an immediate positive jump, the results reveal a generally null or slightly negative short-term impact.

For grandmothers, the LATE estimates for the pooled sample are statistically insignificant and hover around zero (e.g., -0.5 percentage points). However, a statistically significant negative

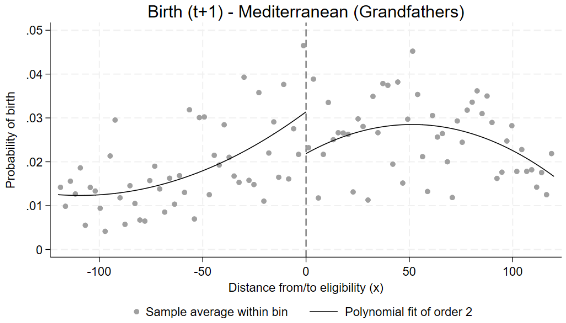


Figure 7: Birth (t+1) - Sons - Mediterranean

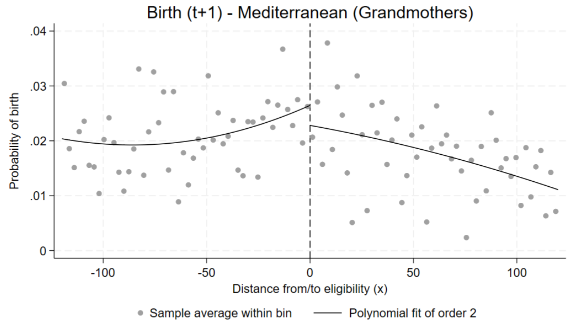


Figure 8: Birth (t+1) - Daughters - Mediterranean

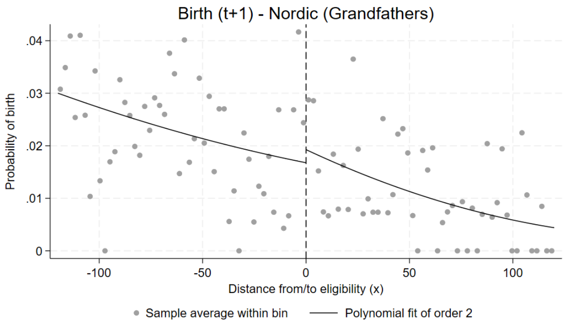


Figure 9: Birth (t+1) - Sons - Nordic

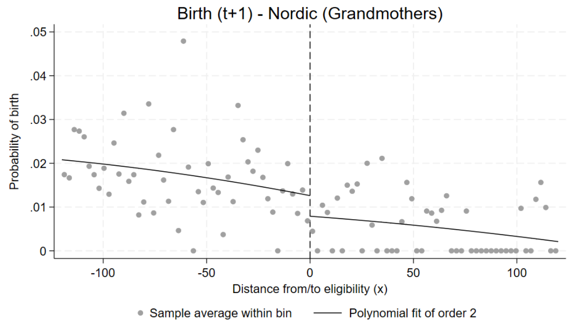


Figure 10: Birth (t+1) - Daughters - Nordic

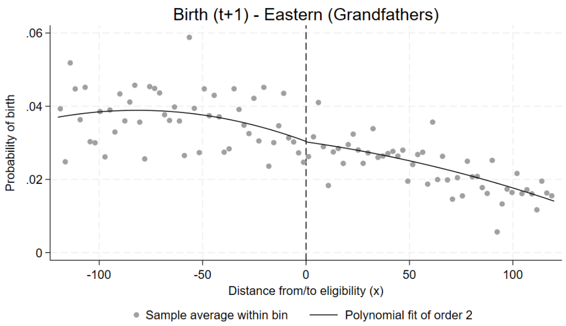


Figure 11: Birth (t+1) - Sons - Eastern

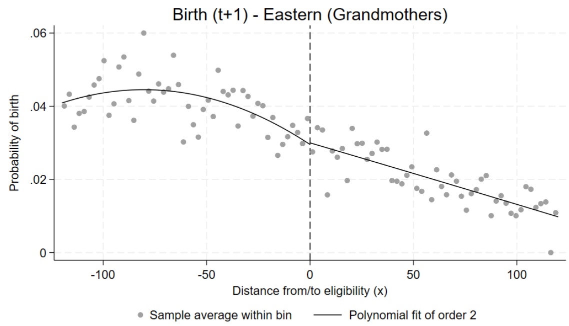


Figure 12: Birth (t+1) - Daughters - Eastern

effect emerges when looking specifically at sons, corresponding to a reduction of 1.2 percentage points. For grandfathers, we actually observe a small but statistically significant negative effect across the board, equal to a reduction of 0.8 percentage points in the probability of childbirth for the pooled sample.

Rather than contradicting the literature, this negative and null tendency is highly informative and perfectly aligns with our theoretical framework regarding short-term rigidities. In the immediate aftermath of retirement, the positive *time effect* (increased availability of informal childcare) has not yet had time to materialize into actual births due to the biological nine-month gestation lag. Conversely, the *negative income effect*—the sudden and often anticipated drop in parental income upon exiting the labor force—acts as an immediate binding constraint. Adult children may temporarily delay fertility decisions to assess the new financial equilibrium of the extended family, particularly in contexts where elderly fathers historically represent the main source of financial support within the extended family..

This dynamic becomes particularly interesting when exploring institutional heterogeneity. The negative short-term impact is mainly driven by the Eastern and Continental regimes, where grandfather retirement significantly reduces childbirth probability by 1.4 and 0.7 percentage points, respectively. Conversely, in the Mediterranean regime, the coefficients are highly noisy and statistically indistinguishable from zero (e.g., +0.9 percentage points for grandfathers and -0.1 percentage points for grandmothers). This is remarkably consistent with the structural characteristics of Southern Europe: while the sudden drop in the grandparents' income acts as a negative shock, it is likely immediately counterbalanced by the extreme necessity of informal childcare in these countries, where formal services are scarce. The tension between the negative income effect and the high value attached to the positive time effect results in a null net impact on fertility timing at $t + 1$.

As highlighted by Bongaarts and Feeney (1998) regarding the distinction between the *tempo* and *quantum* of fertility, short-term adjustments at $t + 1$ could simply reflect a temporary delay. However, exploring the medium-term fertility responses at t+2 and t+3 (Table A.6 and Table A.7) reveals a striking pattern: the negative impact of parental retirement is not a mere short-term rigidity, but rather a persistent and structurally binding constraint.

Across the pooled sample, the causal effect of grandfathers' retirement remains negative and highly significant, reducing childbirth probability by 0.8 percentage points at $t + 2$ and 1.1 percentage points at $t + 3$. The effect for grandmothers also shows a significant reduction of 1.0 percentage points at $t + 2$, before dissipating at $t + 3$.

This persistent negative trajectory provides little support for the hypothesis of a delayed positive *time effect*. Instead, it strongly suggests that the *negative income effect* dominates the medium-term fertility decisions of adult children. The contraction of the extended family's income following retirement reduces the capacity of elderly parents to provide financial transfers to their adult children.

This mechanism is particularly evident by the Mediterranean regime. While the effect at $t + 1$ and $t + 2$ was noisy and counterbalanced by the immediate need for informal childcare, at $t + 3$ the retirement of grandfathers triggers a large and statistically significant reduction in fertility of 4.9 percentage points ($p < 0.01$). This drop is particularly severe for daughters (-6.2 percentage points). In Southern European countries, where formal welfare is weak, young couples heavily rely on the financial support of their parents—historically the male breadwinners—to afford the high costs of child-rearing, housing, and formal childcare. When this financial safety net shrinks due to retirement, the increased availability of grandparental time is simply not enough to compensate for the loss of income, forcing adult children to revise their fertility plans downward. Conversely, in the Nordic regime, where young families are financially independent and supported by strong public childcare infrastructure, parental retirement yields effects that are economically and statistically close to zero across all time horizons.

Table 5: Reduced-Form and Second-Stage Estimates (t+1)

Group of Countries	Birth t+1	Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Reduced-Form	-0.010***	-0.002	-0.003**
		(0.001)	(0.001)	(0.001)
	Second Stage	-0.008***	-0.007	-0.010**
		(0.003)	(0.004)	(0.004)
Continental	Reduced-Form	-0.002**	-0.001	-0.004**
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.007*	-0.003	-0.012**
		(0.004)	(0.005)	(0.005)
Mediterranean	Reduced-Form	0.001	-0.001	0.002
		(0.002)	(0.003)	(0.003)
	Second Stage	0.009	-0.007	0.030
		(0.016)	(0.021)	(0.024)
Nordic	Reduced-Form	-0.001	0.001	-0.003
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.002	0.003	-0.008
		(0.007)	(0.009)	(0.009)
Eastern	Reduced-Form	-0.006***	-0.006*	-0.006*
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.014**	-0.013	-0.015*
		(0.006)	(0.009)	(0.008)
B. Grandmothers				
All Countries	Reduced-Form	-0.001	-0.003*	0.001
		(0.001)	(0.001)	(0.001)
	Second Stage	-0.005	-0.012**	0.002
		(0.004)	(0.006)	(0.006)
Continental	Reduced-Form	-0.001	-0.004**	0.002
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.006	-0.018	0.007
		(0.006)	(0.008)	(0.009)
Mediterranean	Reduced-Form	-0.000	-0.002	0.002
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.001	-0.016	0.016
		(0.013)	(0.018)	(0.021)
Nordic	Reduced-Form	-0.001	-0.004	0.001
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.005	-0.011	0.002
		(0.008)	(0.012)	(0.011)
Eastern	Reduced-Form	-0.001	0.001	-0.004
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.005	0.005	-0.017
		(0.010)	(0.013)	(0.015)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for quadratic age controls for adult children and elderly parents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***\ p < 0.01$, $**\ p < 0.05$, $*\ p < 0.1$.

4.3 Dynamic Responses and Forward-Looking Behavior

A crucial aspect of intergenerational dynamics is the possibility of forward-looking behavior. As highlighted in the recent literature (e.g., Carta and De Philippis (2024)), retirement is rarely an unexpected shock. Adult children might anticipate their parents' transition out of the labor force and strategically adjust their fertility plans before the actual eligibility threshold is reached. In such a scenario, fertility decisions are intrinsically dynamic, and the estimated LATE at the cutoff reflects a combination of immediate constraints and anticipatory responses.

To formally investigate this mechanism, we estimate our baseline Fuzzy RDD models on pre-treatment periods, using childbirths occurring one $(t-1)$, two $(t-2)$, and three $(t-3)$ years prior to parental retirement eligibility as our dependent variables. The reduced-form results, extensively reported in the Appendix for both grandfathers and grandmothers, reveal striking evidence of anticipation effects that are consistent with our theoretical framework based on the *negative income effect*.

First, we observe a highly significant and persistent anticipatory drop in fertility driven by the Eastern regime. In these countries, the impending retirement of either the grandfather or the grandmother significantly reduces childbirth probability by approximately 0.7 to 0.8 percentage points ($p < 0.01$) across all pre-treatment periods ($t-1$, $t-2$, and $t-3$). This indicates that adult children pre-emptively delay childbearing years in advance, bracing for the impending contraction in the extended family's financial resources regardless of which parent is retiring.

The most fascinating dynamic, however, emerges in the Mediterranean regime, where the anticipation behavior exhibits a clear gender asymmetry. As previously discussed, at $t-2$ and $t-3$ prior to the *grandfather's* retirement, we estimate a positive and highly significant increase in fertility of 0.7 percentage points. Conversely, the coefficients prior to the *grandmother's* retirement are highly noisy and statistically indistinguishable from zero across all pre-treatment horizons. This evidence suggests strategic fertility timing by Southern European couples: lacking formal welfare support, young adults accelerate their fertility plans to coincide specifically with the final peak-earning years of the male breadwinner (the grandfather). They may attempt to benefit from the financial support they can receive while the grandfather is still fully employed, knowing that his impending retirement is expected to reduce the family's available financial resources. The impending retirement of the grandmother, whose historical labor income in these cohorts is typically lower, does not trigger the same strategic acceleration.

Far from invalidating our empirical design, these pre-treatment responses demonstrate that European couples behave as forward-looking households. They confirm that the drop in parental income is a economically relevant event that forces adult children to strategically re-time their fertility, either by anticipating births while parental income is still relatively high (Southern Europe) or by pre-emptively delaying them as the income shock approaches (Eastern Europe).

Table 6: Reduced-Form and Second Stage -Estimates (t-1)

Group of Countries	Birth t-1	Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Reduced-Form	-0.003***	-0.002	-0.005**
		(0.001)	(0.001)	(0.002)
	Second Stage	-0.014***	-0.008	-0.021***
		(0.005)	(0.006)	(0.007)
Continental	Reduced-Form	-0.003**	-0.001	-0.005**
		(0.002)	(0.002)	(0.002)
	Second Stage	-0.012*	-0.002	-0.023**
		(0.007)	(0.009)	(0.010)
Mediterranean	Reduced-Form	0.000	-0.001	0.003
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.000	-0.012	0.014
		(0.015)	(0.019)	(0.022)
Nordic	Reduced-Form	-0.007**	-0.005	-0.009*
		(0.003)	(0.005)	(0.005)
	Second Stage	-0.022**	-0.017	-0.027*
		(0.010)	(0.014)	(0.014)
Eastern	Reduced-Form	-0.007**	-0.003*	-0.011**
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.025**	-0.010	-0.043***
		(0.011)	(0.014)	(0.017)
B. Grandmothers				
All Countries	Reduced-Form	-0.002*	-0.001	-0.002
		(0.001)	(0.001)	(0.001)
	Second Stage	-0.008**	-0.007	-0.009*
		(0.003)	(0.005)	(0.005)
Continental	Reduced-Form	-0.000	0.000	-0.001
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.001	-0.001	-0.002
		(0.004)	(0.006)	(0.006)
Mediterranean	Reduced-Form	-0.002	-0.004	0.001
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.016	-0.030	0.003
		(0.016)	(0.021)	(0.024)
Nordic	Reduced-Form	0.000	0.001	-0.001
		(0.003)	(0.004)	(0.004)
	Second Stage	0.000	0.001	-0.002
		(0.008)	(0.011)	(0.011)
Eastern	Reduced-Form	-0.007***	-0.005	-0.010***
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.015**	-0.007	-0.023**
		(0.006)	(0.009)	(0.009)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for potential parents and potential grandparents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$.

5 Conclusions and policy implications

This study was designed to investigate the causal link between grandparental retirement and the fertility decisions of adult children in a pan-European context. Motivated by Europe's persistent demographic challenges—low fertility rates and a progressively ageing population—we aimed to test whether the positive effects identified in single-country studies, such as Germany (Eibich and Siedler, 2020), could be generalized across countries characterized by different welfare systems and family support structures. By employing a fuzzy Regression Discontinuity Design on SHARE data, we exploited the exogenous variation in retirement status induced by statutory pension eligibility rules across four distinct welfare regimes to provide robust causal estimates.

Our findings reveal a nuanced and theoretically rich pattern that challenges the traditional view of retirement as a purely positive shock to informal childcare availability. We confirm the existence of a strong and highly significant first stage across all welfare regimes, demonstrating that pension eligibility acts as a powerful shock to labor supply. However, the second-stage results show that the immediate causal impact of parental retirement on adult children's fertility at $t+1$ is generally null or slightly negative.

Crucially, the medium-term and pre-treatment dynamics suggest that the delayed positive *time effect* does not dominate fertility decisions. As documented for $t+2$ and $t+3$, the negative impact of parental retirement persists over time, emerging as a persistent constraint on fertility choices. The evidence therefore suggests that the *negative income effect*—the reduction in parental income following retirement—plays a central role in shaping fertility outcomes (Aparicio-Fenoll and Vidal-Fernandez, 2015). This mechanism is particularly evident in the Mediterranean regime, where the retirement of the grandfather is associated with a substantial decline in fertility at $t+3$.

Furthermore, our analysis of pre-treatment periods highlights important forward-looking behaviour. Adult children appear to strategically adjust fertility timing in anticipation of parental retirement and the associated changes in family resources. In Southern Europe, couples tend to anticipate childbirths while parental income is still relatively high, whereas in Eastern Europe fertility appears to be postponed in anticipation of the future reduction in the extended family's financial resources.

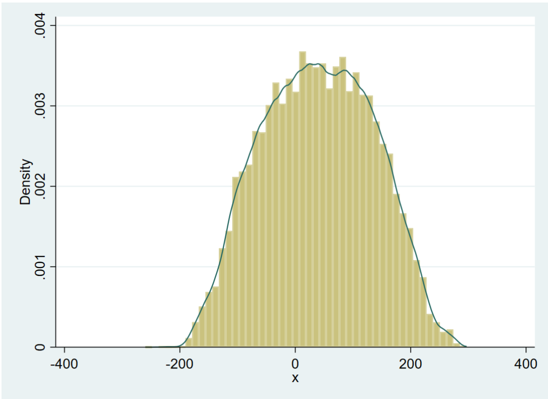
From a policy perspective, our findings suggest the need for a reconsideration of the demographic side effects of pension reforms. Public debate often assumes that raising the retirement age affects fertility mainly through the reduction in grandparental childcare time. Our results instead point to a broader structural vulnerability: the strong financial dependence of young couples on the labor income of the older generation. In welfare regimes characterized by limited formal childcare provision and weak family support policies, young adults heavily rely on parental earnings to sustain child-rearing costs, housing expenditures, and broader family consumption. Consequently, pension reforms that reduce the income of older generations may generate unintended negative effects on fertility decisions.

To mitigate ongoing demographic decline, policymakers cannot rely exclusively on the indirect channel of grandparental retirement. Instead, they should implement direct structural interventions aimed at strengthening the economic independence of younger generations. Policies such as affordable formal childcare, family-friendly labor market arrangements, and targeted financial support for families may reduce the dependence of fertility decisions on parental employment and retirement dynamics.

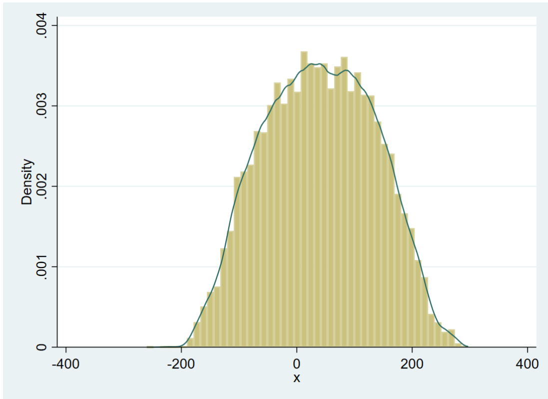
Appendix

A Additional descriptive

Figure A.1: RD Plot: Kernel Density

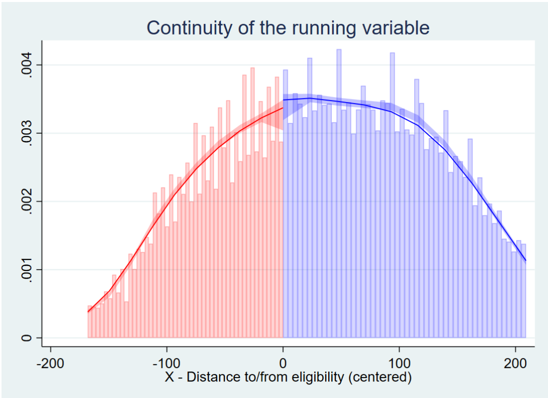


(**a**) Grandfathers

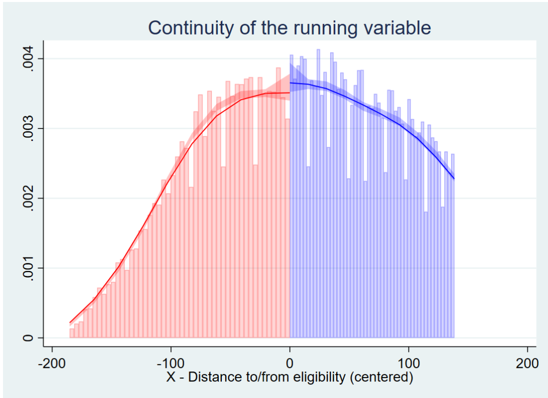


(**b**) Grandmothers

Figure A.2: Manipulation test



(**a**) Grandfathers



(**b**) Grandmothers

Table A.1: Descriptive statistics: Outcome variables and covariates for adult children by Country

Country	birth	birth (t+1)	birth (t+2)	birth (t+3)	birth within 2 years	birth within 3 years	n of children	gender	age	distance from parents	marital status	contacts with parents	partner	employment status	education
Austria	0.0157	0.0077	0.0056	0.0019	0.0205	0.0481	0.8200	0.5087	36.0177	4.4959	2.7313	3.4812	2.5682	2.6329	7.3705
	(0.1244)	(0.0878)	(0.0744)	(0.0437)	(0.1418)	(0.2142)	(1.0254)	(0.4999)	(7.5523)	(1.9918)	(1.5021)	(1.9415)	(1.5136)	(8.081)	(14.7375)
Germany	0.01778	0.0104	0.0038	0.0006	0.0254	0.0552	0.7731	0.5121	34.8871	4.7545	2.8706	3.4486	2.7836	3.5007	6.8349
	(0.1321)	(0.1015)	(0.0621)	(0.0255)	(0.1575)	(0.2285)	(1.0610)	(0.4998)	(7.6347)	(2.2126)	(1.4768)	(1.9492)	(1.6494)	(10.6995)	(11.4795)
Netherlands	0.0156	0.0076	0.00593	0.0021	0.0247	0.0912	0.8421	0.5166	35.2656	4.5152	2.4831	4.4493	2.6552	3.5104	7.5989
	(0.1240)	(0.0872)	(0.0768)	(0.0458)	(0.1552)	(0.2882)	(1.1202)	(0.4997)	(7.7878)	(1.9555)	(1.4161)	(1.3786)	(1.4623)	(11.4870)	(13.1639)
Spain	0.0175	0.0103	0.0064	0.0043	0.0272	0.0616	0.8532	0.5263	36.0490	3.8108	2.3940	4.3821	1.8812	2.7491	12.9120
	(0.1311)	(0.1011)	(0.0799)	(0.0661)	(0.1628)	(0.2406)	(1.1614)	(0.4993)	(7.4816)	(2.2411)	(1.4552)	(1.4459)	(1.3262)	(7.8700)	(27.0057)
Italy	0.0139	0.0100	0.0076	0.0040	0.0263	0.0575	0.7849	0.5254	36.3484	3.4942	2.4208	4.7534	1.6499	5.3654	9.6416
	(0.1174)	(0.0996)	(0.0873)	(0.0628)	(0.1600)	(0.2329)	(1.1165)	(0.4993)	(7.5385)	(2.2623)	(1.4628)	(0.9621)	(1.1595)	(16.2251)	(14.2904)
Denmark	0.0228	0.01230	0.0075	0.0026	0.0295	0.0591	0.9744	0.5067	33.2675	5.0421	2.9398	3.3237	2.7836	3.9747	7.3484
	(0.1494)	(0.1102)	(0.0865)	(0.0513)	(0.1693)	(0.2359)	(1.1459)	(0.4997)	(8.0738)	(1.8524)	(1.4726)	(1.9739)	(1.4997)	(12.1283)	(12.4879)
Switzerland	0.0192	0.0118	0.0090	0.0055	0.0292	0.0601	0.8127	0.5063	34.5729	4.4905	2.7035	3.7464	2.8704	3.8398	6.9958
	(0.1373)	(0.1081)	(0.0947)	(0.0745)	(0.1684)	(0.2378)	(1.0829)	(0.4999)	(7.7940)	(2.0062)	(1.4971)	(1.8558)	(1.5589)	(12.0981)	(15.4889)
Belgium	0.0179	0.0011	0.0074	0.0032	0.0284	0.0592	0.9467	0.5118	34.6229	4.2287	2.5150	4.3690	2.6454	4.8668	13.8794
	(0.1328)	(0.1083)	(0.0857)	(0.0568)	(0.1661)	(0.2361)	(1.1589)	(0.4998)	(7.7437)	(2.0186)	(1.3648)	(1.4581)	(1.6661)	(16.4082)	(13.8901)
Poland	0.0112	0.0051	0.0031	0	0.0146	0.0286	1.1195	0.5079	36.4973	4.2203	2.2406	3.9272	2.2841	3.8238	13.4517
	(0.1056)	(0.0717)	(0.0562)	(0)	(0.1202)	(0.1669)	(1.1468)	(0.4999)	(7.0710)	(2.3662)	(1.5363)	(1.7723)	(1.4958)	(13.3944)	(18.4858)
Luxembourg	0.0208	0.0051	0	0	0.0123	0.0466	0.7493	0.5237	33.0162	4.3093	2.8547	3.9714	2.3159	4.0334	15.5022
	(0.1429)	(0.0715)	(0)	(0)	(0.1105)	(0.2113)	(1.1555)	(0.4995)	(7.3338)	(2.2493)	(1.4192)	(1.7433)	(1.5043)	(13.7297)	(17.0344)
Hungary	0.0190	0.0051	0	.	0.0229	1	0.7884	0.5081	34.9620	4.0141	2.5583	4.4699	1.9078	3.2756	14.954
	(0.1365)	(0.0718)	(0)	(.)	(0.1500)	(0)	(1.0783)	(0.5000)	(7.2362)	(2.4140)	(1.4817)	(1.3570)	(1.3916)	(11.5182)	(28.6294)
Slovenia	0.0262	0.0133	0.0051	0.0028	0.0328	0.0867	1.1144	0.5091	35.9072	3.6982	2.2755	4.7346	1.8475	7.226	9.6013
	(0.1597)	(0.1146)	(0.0711)	(0.0536)	(0.1783)	(0.2815)	(1.1194)	(0.4999)	(7.1269)	(2.0907)	(1.2677)	(0.9957)	(1.2525)	(22.4063)	(16.4241)

B Additional results

B.1 First stage

Table A.2: First stage by groups of country (t+2)

Group of Countries		Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Eligibility	0.231***	0.231***	0.232***
	dummy (Z)	(0.003)	(0.004)	(0.004)
Continental	Eligibility	0.220***	0.218***	0.221***
	dummy (Z)	(0.004)	(0.006)	(0.006)
Mediterranean	Eligibility	0.142***	0.143***	0.141***
	dummy (Z)	(0.007)	(0.010)	(0.010)
Nordic	Eligibility	0.342***	0.334***	0.349***
	dummy (Z)	(0.009)	(0.012)	(0.012)
Eastern	Eligibility	0.284***	0.297***	0.271***
	dummy (Z)	(0.008)	(0.011)	(0.011)
B. Grandmothers				
All Countries	Eligibility	0.289***	0.285***	0.294***
	dummy (Z)	(0.003)	(0.004)	(0.004)
Continental	Eligibility	0.286***	0.280***	0.292***
	dummy (Z)	(0.004)	(0.006)	(0.006)
Mediterranean	Eligibility	0.120***	0.126***	0.114***
	dummy (Z)	(0.007)	(0.010)	(0.010)
Nordic	Eligibility	0.379***	0.380***	0.378***
	dummy (Z)	(0.009)	(0.012)	(0.012)
Eastern	Eligibility	0.341***	0.332***	0.350***
	dummy (Z)	(0.005)	(0.008)	(0.008)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for potential parents and potential grandparents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$.

Table A.3: First stage by groups of country (t+3)

Group of Countries		Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Eligibility	0.231***	0.231***	0.232***
	dummy (Z)	(0.003)	(0.004)	(0.004)
Continental	Eligibility	0.220***	0.218***	0.221***
	dummy (Z)	(0.004)	(0.006)	(0.006)
Mediterranean	Eligibility	0.142***	0.143***	0.141***
	dummy (Z)	(0.007)	(0.010)	(0.010)
Nordic	Eligibility	0.342***	0.334***	0.349***
	dummy (Z)	(0.009)	(0.012)	(0.012)
Eastern	Eligibility	0.284***	0.297***	0.271***
	dummy (Z)	(0.008)	(0.011)	(0.011)
B. Grandmothers				
All Countries	Eligibility	0.289***	0.285***	0.294***
	dummy (Z)	(0.003)	(0.004)	(0.004)
Continental	Eligibility	0.286***	0.280***	0.292***
	dummy (Z)	(0.004)	(0.006)	(0.006)
Mediterranean	Eligibility	0.120***	0.126***	0.114***
	dummy (Z)	(0.007)	(0.010)	(0.010)
Nordic	Eligibility	0.379***	0.380***	0.378***
	dummy (Z)	(0.009)	(0.012)	(0.012)
Eastern	Eligibility	0.341***	0.332***	0.350***
	dummy (Z)	(0.005)	(0.008)	(0.008)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for potential parents and potential grandparents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$.

Table A.4: Reduced-Form and Second Stage -Estimates (t-2)

Group of Countries	Birth t-2	Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Reduced-Form	-0.001	-0.001	-0.003
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.004	0.003	-0.011
		(0.005)	(0.007)	(0.008)
Continental	Reduced-Form	-0.003	0.001	-0.006**
		(0.002)	(0.002)	(0.002)
	Second Stage	-0.010	0.005	-0.025***
		(0.008)	(0.010)	(0.011)
Mediterranean	Reduced-Form	0.007***	0.004	-0.009***
		(0.002)	(0.003)	(0.003)
	Second Stage	0.036**	0.021	0.052**
		(0.015)	(0.021)	(0.022)
Nordic	Reduced-Form	-0.006	-0.002	-0.009*
		(0.004)	(0.005)	(0.005)
	Second Stage	-0.019*	-0.009	-0.029*
		(0.011)	(0.016)	(0.016)
Eastern	Reduced-Form	-0.004	-0.002	-0.005
		(0.003)	(0.004)	(0.005)
	Second Stage	-0.012	-0.006	-0.018
		(0.012)	(0.015)	(0.018)
B. Grandmothers				
All Countries	Reduced-Form	-0.001	-0.001	-0.001
		(0.001)	(0.001)	(0.002)
	Second Stage	-0.006	-0.007	-0.004
		(0.004)	(0.005)	(0.005)
Continental	Reduced-Form	0.001	-0.001	0.002
		(0.001)	(0.002)	(0.002)
	Second Stage	0.001	-0.005	0.007
		(0.004)	(0.006)	(0.007)
Mediterranean	Reduced-Form	-0.002	-0.004	0.001
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.025	-0.038	-0.009
		(0.017)	(0.023)	(0.026)
Nordic	Reduced-Form	0.004	0.005	0.002
		(0.003)	(0.004)	(0.005)
	Second Stage	0.010	0.013	0.006
		(0.009)	(0.012)	(0.013)
Eastern	Reduced-Form	-0.008***	-0.005	-0.011***
		(0.002)	(0.003)	(0.004)
	Second Stage	-0.015**	-0.006	-0.025***
		(0.007)	(0.009)	(0.010)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for potential parents and potential grandparents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$.

Table A.5: Reduced-Form and Second Stage -Estimates (t-3)

Group of Countries	Birth t-3	Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Reduced-Form	0.001	0.003*	-0.000
		(0.001)	(0.002)	(0.002)
	Second Stage	0.008	0.013*	0.003
		(0.006)	(0.008)	(0.009)
Continental	Reduced-Form	0.000	0.003	-0.003
		(0.002)	(0.002)	(0.003)
	Second Stage	0.005	0.018	-0.009
		(0.009)	(0.011)	(0.013)
Mediterranean	Reduced-Form	0.007***	0.005	0.008**
		(0.002)	(0.003)	(0.003)
	Second Stage	0.033**	0.027	0.039*
		(0.016)	(0.022)	(0.023)
Nordic	Reduced-Form	-0.003	0.002	-0.006
		(0.004)	(0.005)	(0.005)
	Second Stage	-0.009	-0.004	-0.023
		(0.013)	(0.018)	(0.018)
Eastern	Reduced-Form	-0.000	-0.001	0.000
		(0.003)	(0.005)	(0.005)
	Second Stage	0.003	-0.002	0.009
		(0.012)	(0.016)	(0.019)
B. Grandmothers				
All Countries	Reduced-Form	-0.001	-0.001	-0.000
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.006	-0.007	-0.004
		(0.004)	(0.005)	(0.006)
Continental	Reduced-Form	0.000	-0.000	0.001
		(0.002)	(0.002)	(0.002)
	Second Stage	-0.000	-0.004	0.004
		(0.005)	(0.007)	(0.007)
Mediterranean	Reduced-Form	-0.001	-0.003	0.002
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.021	-0.025	-0.016
		(0.018)	(0.025)	(0.027)
Nordic	Reduced-Form	0.000	0.001	0.001
		(0.003)	(0.005)	(0.005)
	Second Stage	-0.001	-0.006	0.003
		(0.010)	(0.014)	(0.015)
Eastern	Reduced-Form	-0.007**	-0.005	-0.009**
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.008	-0.000	-0.016
		(0.007)	(0.010)	(0.011)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for potential parents and potential grandparents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$.

Table A.6: Reduced-Form and Second Stage -Estimates (t+2)

Group of Countries	**Birth t+2**	**Adult children - all**	**Sons**	**Daughters**
A. Grandfathers				
All Countries	Reduced-Form	-0.003***	-0.002	-0.004***
		(0.001)	(0.001)	(0.001)
	Second Stage	-0.008***	-0.007	-0.010***
		(0.003)	(0.004)	(0.004)
Continental	Reduced-Form	-0.003**	-0.002	-0.004**
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.007**	-0.003	-0.012***
		(0.004)	(0.005)	(0.005)
Mediterranean	Reduced-Form	-0.003	-0.002	-0.005
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.009	-0.007	0.030
		(0.016)	(0.021)	(0.024)
Nordic	Reduced-Form	-0.001	0.001	-0.004
		(0.002)	(0.003)	(0.004)
	Second Stage	-0.002	0.003	-0.008
		(0.007)	(0.009)	(0.009)
Eastern	Reduced-Form	-0.004*	-0.004	-0.004
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.014**	-0.013	-0.015*
		(0.006)	(0.009)	(0.008)
B. Grandmothers				
All Countries	Reduced-Form	-0.002**	-0.001	-0.004**
		(0.001)	(0.001)	(0.001)
	Second Stage	-0.010***	-0.007	-0.012***
		(0.003)	(0.004)	(0.004)
Continental	Reduced-Form	-0.003**	-0.003	-0.003
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.010***	-0.006	-0.015***
		(0.004)	(0.005)	(0.005)
Mediterranean	Reduced-Form	-0.003	-0.003	-0.002
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.020	-0.017	-0.021
		(0.017)	(0.022)	(0.026)
Nordic	Reduced-Form	-0.001	0.002	-0.003
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.003	0.003	-0.011
		(0.006)	(0.009)	(0.009)
Eastern	Reduced-Form	-0.001	0.006	-0.006
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.009	-0.008	-0.010
		(0.006)	(0.009)	(0.008)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for potential parents and potential grandparents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$.

Table A.7: Reduced-Form and Second Stage -Estimates (t+3)

Group of Countries	Birth t+3	Adult children - all	Sons	Daughters
A. Grandfathers				
All Countries	Reduced-Form	-0.003***	-0.002*	-0.004***
		(0.001)	(0.001)	(0.001)
	Second Stage	-0.011***	-0.008*	-0.014***
		(0.003)	(0.004)	(0.004)
Continental	Reduced-Form	-0.002	-0.000	-0.004**
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.007*	-0.002	-0.012**
		(0.004)	(0.005)	(0.005)
Mediterranean	Reduced-Form	-0.006**	-0.004	-0.008**
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.049***	-0.035	-0.062**
		(0.018)	(0.023)	(0.029)
Nordic	Reduced-Form	-0.004	-0.002	-0.006*
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.009	-0.005	-0.015*
		(0.006)	(0.008)	(0.008)
Eastern	Reduced-Form	-0.004**	-0.006*	-0.003
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.011*	-0.014	-0.009
		(0.006)	(0.009)	(0.008)
B. Grandmothers				
All Countries	Reduced-Form	-0.002	-0.001	-0.002
		(0.001)	(0.001)	(0.001)
	Second Stage	-0.007	-0.004	-0.009
		(0.004)	(0.006)	(0.006)
Continental	Reduced-Form	-0.001	-0.001	-0.001
		(0.001)	(0.002)	(0.002)
	Second Stage	-0.006	-0.006	-0.006
		(0.006)	(0.008)	(0.009)
Mediterranean	Reduced-Form	-0.003	-0.001	-0.004
		(0.002)	(0.003)	(0.003)
	Second Stage	-0.017	-0.008	-0.026
		(0.014)	(0.017)	(0.021)
Nordic	Reduced-Form	0.000	0.000	0.000
		(0.003)	(0.004)	(0.004)
	Second Stage	0.001	0.001	0.001
		(0.007)	(0.010)	(0.009)
Eastern	Reduced-Form	-0.000	0.001	-0.002
		(0.003)	(0.004)	(0.004)
	Second Stage	-0.013	-0.009	-0.015
		(0.010)	(0.014)	(0.014)

Note: Standard errors, clustered at the adult child level, are shown in parentheses. All estimations include fixed effects, along with quadratic age terms for potential parents and potential grandparents. The age of adult children ranges from 20 to 50 years, and the age of elderly parents ranges from 40 to 80 years. Significance levels: $***$ $p < 0.01$, $**$ $p < 0.05$, $*$ $p < 0.1$.

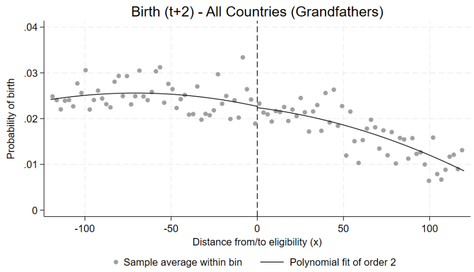


Figure B.1: Birth (t+2) - Sons - All country

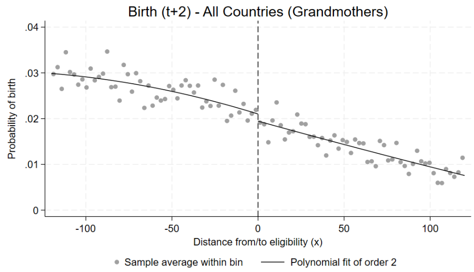


Figure B.2: Birth (t+2) - Daughters - All country

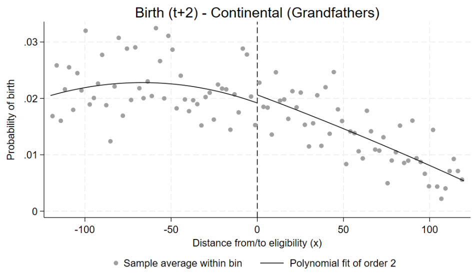


Figure B.3: Birth (t+2) - Sons - Continental

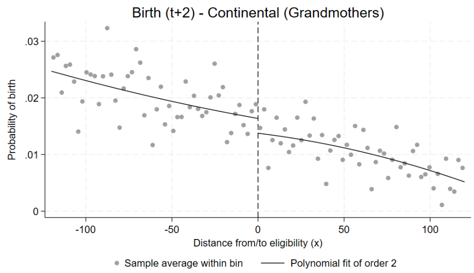


Figure B.4: Birth (t+2) - Daughters - Continental

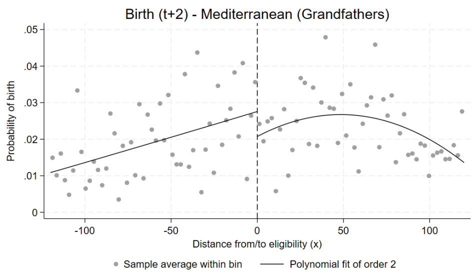


Figure B.5: Birth (t+2) - Sons - Mediterranean

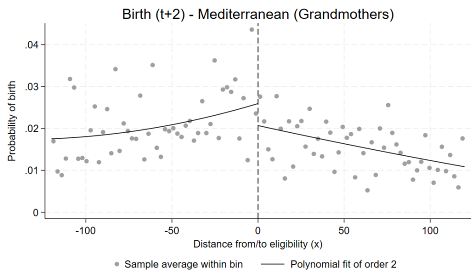


Figure B.6: Birth (t+2) - Daughters - Mediterranean

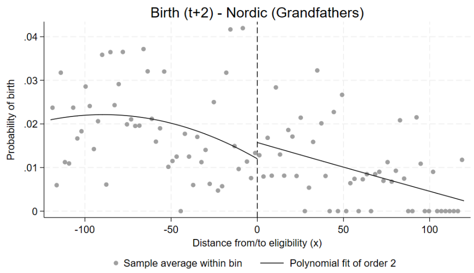


Figure B.7: Birth (t+2) - Sons - Nordic

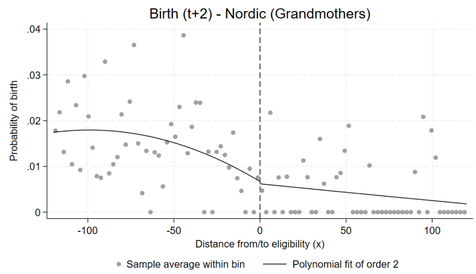


Figure B.8: Birth (t+2) - Daughters - Nordic

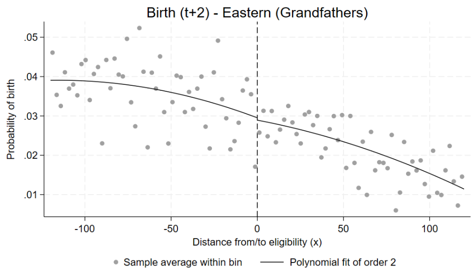


Figure B.9: Birth (t+2) - Sons - Eastern

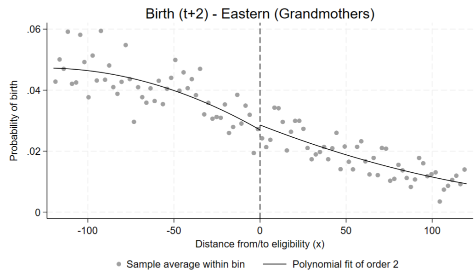


Figure B.10: Birth (t+2) - Daughters - Eastern

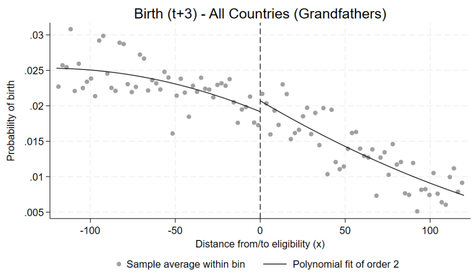


Figure B.11: Birth (t+3) - Sons - All country

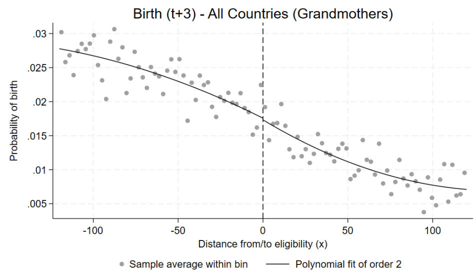


Figure B.12: Birth (t+3) - Daughters - All country

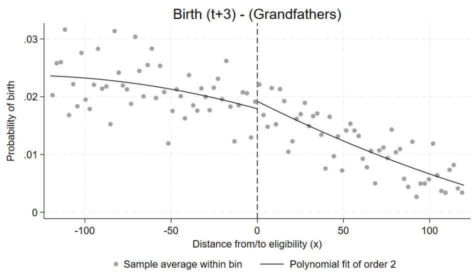


Figure B.13: Birth (t+3) - Sons - Continental

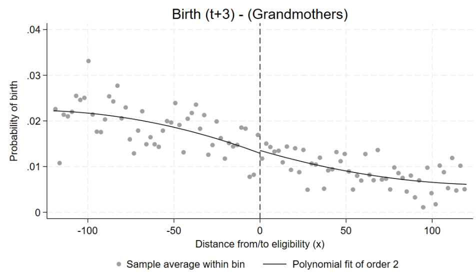


Figure B.14: Birth (t+3) - Daughters - Continental

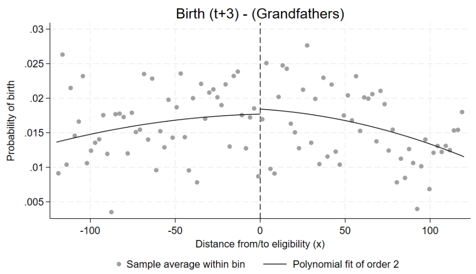


Figure B.15: Birth (t+3) - Sons - Mediterranean

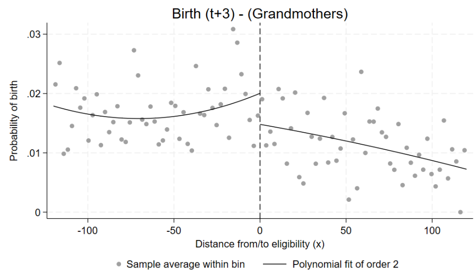


Figure B.16: Birth (t+3) - Daughters - Mediterranean

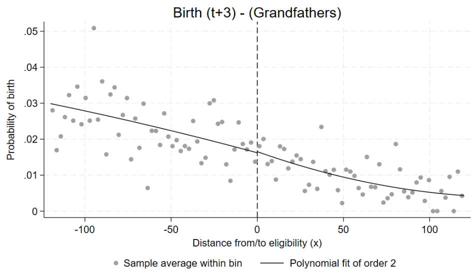


Figure B.17: Birth (t+3) - Sons - Nordic

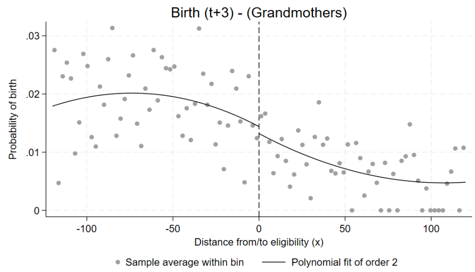


Figure B.18: Birth (t+3) - Daughters - Nordic

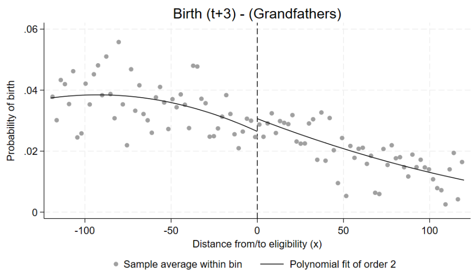


Figure B.19: Birth (t+3) - Sons - Eastern

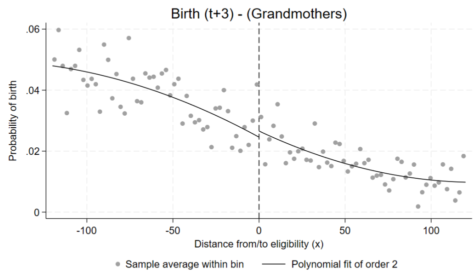


Figure B.20: Birth (t+3) - Daughters - Eastern

Bibliography

Akyol, P. and Atalay, K. (2025). The intergenerational impact of pension reforms: How grandmothers' pension eligibility affects daughters' fertility. *Economics Letters*, 248:112239.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Aparicio-Fenoll, A. and Vidal-Fernandez, M. (2015). Working women and fertility: the role of grandmothers' labor force participation. *CESifo Economic Studies*, 61(1):123–147.

Backhaus, A. and Barslund, M. (2021). The effect of grandchildren on grandparental labor supply: Evidence from europe. *European Economic Review*, 137:103817.

Baldi, M. M., Coppier, R., and Michetti, E. (2025). Elderly labor supply, endogenous grandparental childcare, and fertility in an olg model. *Economic Modelling*, page 107102.

Bongaarts, J. and Feeney, G. (1998). On the quantum and tempo of fertility. *Population and development review*, pages 271–291.

Cai, T., Ma, S., and Zhong, R. (2024). Does parents' retirement influence the fertility intentions of subsequent generations among internal migrant populations? evidence from china. *Plos one*, 19(10):e0311028.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, 101(3):442–451.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2015). Optimal data-driven regression discontinuity plots. *Journal of the American Statistical Association*, 110(512):1753–1769.

Carta, F. and De Philippis, M. (2024). The forward-looking effect of increasing the full retirement age. *The Economic Journal*, 134(657):165–192.

Cavallini, F. (2024). Not the right time for children: Unemployment, fertility, and abortion. *Labour Economics*, 87:102492.

Eibich, P. and Siedler, T. (2020). Retirement, intergenerational time transfers, and fertility. *European Economic Review*, 124:103392.

Gelman, A. and Imbens, G. (2019). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Ilciukas, J. (2023). Fertility and parental retirement. *Journal of Public Economics*, 226:104928.

Malisa, A. (2024). Grand-parenthood and retirement. *Labour Economics*, 90:102547.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Porter (2003). Estimation in the regression discontinuity model. unpublished manuscript. *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison, p 5-19.*

Rupert, P. and Zanella, G. (2018). Grandchildren and their grandparents' labor supply. *Journal of Public Economics*, 159:89–103.

Salvati, L., Benassi, F., Miccoli, S., Rabiei-Dastjerdi, H., and Matthews, S. A. (2020). Spatial variability of total fertility rate and crude birth rate in a low-fertility country: Patterns and trends in regional and local scale heterogeneity across italy, 2002–2018. *Applied Geography*, 124:102321.

SHARE-ERIC (2024a). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 2. Release version: 9.0.0.

SHARE-ERIC (2024b). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 4. Release version: 9.0.0.

SHARE-ERIC (2024c). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5. Release version: 9.0.0.

SHARE-ERIC (2024d). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 6. Release version: 9.0.0.

SHARE-ERIC (2024e). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 7. Release version: 9.0.0.

SHARE-ERIC (2024f). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 8. Release version: 9.0.0.

SHARE-ERIC (2024g). Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 9. Release version: 9.0.0.

Stella, L. (2017). Living arrangements in europe: Whether and why paternal retirement matters. *Review of Economics of the Household*, 15(2):497–525.

Zanardo, V., Polver, S., Zanardo, B., Volpe, F., Battistin, T., Ghiro, L., and Straface, G. (2025). Gestational weight gain and postpartum depressive symptoms in high-income women with low and late fertility. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, page 114746.