London School of Economics and Political Science

Essays in Applied Economics

London, September 2019

Giulia Bovini

A thesis submitted to the Department of Economics of the London School of Economics for the degree of Doctor of Philosophy.

Declaration

I certify that the thesis I have presented for examination for the PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I declare that my thesis consists of approximately 75,000 words.

Statement of conjoint work

I confirm that Chapter 1 was jointly co-authored with Matteo Paradisi and I contributed 50% of this work.

I confirm that Chapter 3 was jointly co-authored with Andrés Barrios Fernández and I contributed 50% of this work.

Statement of use of third party for editorial help

I can confirm that my thesis was copy edited for conventions of language, spelling and grammar by the LSE Language Centre.

Acknowledgments

I would like to express my profound gratitude to my supervisor Steve Pischke, for the excellent supervision, caring guidance and continued support throughout the PhD journey. I have benefitted from fruitful interactions with the faculty of the LSE Economics Department, in particular with the members of the CEP, the Labour Economics group, the Public Economics group and the Economics of Education group. I thank them all.

I gratefully acknowledge financial support from the UniCredit Foundation and I am thankful to Annalisa Aleati and Giannantonio De Roni for their encouragement. I am also grateful for the generous financial support provided by the LSE Economics Department.

Parts of the analyses presented in this dissertation have been carried out at the Italian Social Security Institute (INPS), thanks to the VisitINPS Initiative. I am thankful to all those who have been involved in setting up, sponsoring and managing the visiting program. In particular, I thank Tito Boeri, Pietro Garibaldi, Massimo Antichi, Maria Domenica Carnevale, Maria Cozzolino, Edoardo Di Porto, and Paolo Naticchioni. I am also thankful to Massimo Ascione, Elio Bellucci, Luca Cammarata, Vito La Monica and Marcella Nunzi, as well as to all the staff of Direzione Centrale Studi e Ricerche.

I thank my co-authors Andrés Barrios Fernández and Matteo Paradisi, for having shared their knowledge and intellectual curiosity with me.

I would like to thank Paolo Sestito, Matteo Bugamelli and Eliana Viviano at the Bank of Italy, for encouraging me to continue and finish this journey. I am also thankful for the support received from all my colleagues in the Households and Labour Market Division, as well as from many friends in other divisions. I declare that the views and opinions expressed in this dissertation are mine and do not necessarily reflect those of the Bank of Italy.

In London I met several new friends and I am glad to have shared this journey with them. I thank, among others, Alexia, Andrés, Felix, Francesco, Gianpaolo, Giulia, Giuseppe, Inna, Marta, Matteo, Monica, Nicola, and Vincenzo.

I thank Francesco for always standing by my side, in good and difficult times. Without the unconditional support and love of my parents and grandparents, I would not have completed this journey. This dissertation is dedicated to them.

Abstract

The first chapter examines how an increase in statutory retirement ages affects in the short-run firms' labour demand for older workers on the cusp of retirement, their co-workers and outsiders. By leveraging administrative data and a pension reform that in 2011 suddenly and heterogeneously raised statutory retirement ages in Italy, the paper finds effects on firing and hiring that are consistent with substitutability among different types of workers. It also documents changes in the take-up of other social insurance programs. This caused leakages on the revenues raised on a firm's employees by the reform.

The second chapter studies whether pension reforms affect labour supply before individuals come close to the retirement and benefit claiming decisions. By relying on administrative data, the paper examines the labour supply dynamics of a sample of Italian private-sector workers who were aged 35-45 in 1995, when a pension reform changed the way of calculating their social security benefits. Workers are observed until they reach ages between 46 and 56: during this period, at most small adjustments to labour supply are documented. Heterogeneity analyses explore whether the extent of the adjustments differs across socio-demographic groups.

The third chapter assesses the effects of a reform that substantially lengthened the school day in Chilean publicly subsidized primary schools. By leveraging administrative data, the paper documents a positive effect on reading scores in fourth grade. Achievement improves more for schoolchildren who enrolled in no-fee charter schools than in public schools and there is a difference in how these schools adjusted the teaching input; the former, which enjoy more autonomy, relied more than the latter on hiring new classroom teachers and less on increasing working hours per teacher. Albeit less robust, there is also suggestive evidence that positive effects are larger for less advantaged pupils.

Contents

List of Figures

List of Tables

Chapter 1

Labour substitutability and the impact of raising the retirement age

1.1 Introduction

Substitutability among workers within the firm is crucial for the study of public policies. When a fraction of workers are targeted by a policy that alters incentives to work, spillovers on their co-workers may arise depending on whether the employer finds close substitutes or complements among incumbent employees. Labour demand for outsiders can also be affected. These "hidden" micro-level effects have welfare implications that are hard to detect when looking at targeted individuals. Labour substitutability may exist along several worker's characteristics: an example is substitutability between age cohorts that can affect the incidence of Social Security policies targeting older employees. In this paper, we propose a novel approach that regards firms as active agents in the analysis of delayed retirement policies and extends to other policies that lower the turnover of older workers.

Raising the retirement age provides older employees with incentives to postpone retirement.[1] A rich literature documents the positive effects of these reforms on old-age labour supply (*e.g.*, Mastrobuoni, 2009; Staubli and Zweimüller, 2013).[2] There is, on the other hand, limited evidence on how firms adjust labour demand when older employees work for longer than expected. Do firms change their hiring and firing policies? Do labour demand adjustments affect earnings of incumbent employees? Do these changes uniformly affect all cohorts of workers or disproportionately affect some of them? Answering these questions sheds light on the degree of substitutability among workers of different ages within the firm. The extent to which

[1]Many governments have passed reforms that increase the retirement age to cope with population aging and its threat to the sustainability of Social Security systems. The US is following such trends by committing to gradually adjust the full retirement age from 66 to 67 by 2022. The Congressional Budget Office has suggested further raising the statutory age to 70 to help reduce the budget deficit between 2017 and 2026 (Congressional Budget Office, 2016). Most European countries have implemented similar measures since 2000 (Carone et al., 2016).

[2]Other papers on how public pension systems affect retirement behaviour are Behaghel and Blau (2012), Cribb et al. (2016), Lalive et al. (2017), Manoli and Weber (2018), Fetter and Lockwood (2018) and Seibold (2019). This literature documents bunching in retirement at statutory retirement ages and large responses of retirement choices to statutory retirement ages.

older employees close to retirement, their co-workers, and outsiders are substitutes has important implications for the effects of delayed retirement policies as well as of other policies that similarly reduce the turnover of older workers. Therefore, it is primarily important to assess firm responses to changes in the design of Social Security. In addition, since revenues raised on a firm's employees by an increase in the statutory retirement age are a key component of welfare, it is crucial to examine whether and how they are affected by labour substitutability within the firm.

To address these questions we exploit the quasi-experimental variation of a unique pension reform implemented in Italy in 2011 - the *Fornero* reform - that caused sudden, substantial and heterogeneous changes in the full retirement age, *i.e.* the age at which workers can claim full pension benefits.[3] We link matched employer-employee records for small and medium firms to novel records that track all contributions to Social Security accumulated by their workers (more than 6 million individuals, over one third of private employees in 2009). Leveraging this novel match of different data sources, we can build firm-level measures of the reform-induced shock to the retirement date of older employees and we can study whether it affects firms' firing and hiring decisions. Furthermore, we can observe the entire career of all workers and their take-up of various social insurance programs.

We start by investigating whether and how employers change their demand for labour in response to the reform. The results are informative about the degree of substitutability between workers within and across age cohorts. We then document how a higher full retirement age affects labour earnings and the take-up of social insurance programs of older employees and - due to labour substitutability within the firm - of their co-workers. We conclude by discussing the implications of micro-level substitutability for the revenues collected by the reform from a firm's employees in the short-run. To this end, we incorporate the demand-driven behavioural responses into the estimation of the fiscal externality of the policy, *i.e.* the leakages on the revenues that the government hopes to raise on a firm's employees. We thus improve upon the existing literature that focuses on the behavioural responses of older workers. This exercise represents a first step to evaluating the importance of labour substitutability and firms' decisions for the analysis of Social Security policies.

Estimating labour demand responses to reforms that increase the full retirement age poses two main identification challenges. First, most pension reforms are anticipated. Confounding anticipation effects make it hard to isolate firms responses. Second, the share of employees whose retirement date is delayed in the short-run, which is an intuitive metric of the reform shock, depends on the workforce age distribution. Workers' demographics vary across firms possibly due to differences in

[3] The full retirement age is intended throughout the paper as the age at which workers can claim full pension benefits, as opposed to the early retirement age when a different - typically less generous - rule for computing retirement benefits applies.

labour demand trends and other unobservable, time-varying characteristics. Hence, firms with a high concentration of younger workers, which are not affected much by the reform, may not serve as credible controls.

The features of the 2011 Italian pension reform allow us to address both identification issues. The reform was enacted by a newly appointed technocratic government in December 2011 and entered into effect in January 2012, leaving limited room for anticipatory effects. The new law raised the age and contribution requirements for old-age and seniority pensions. The design of the policy generated heterogeneous changes in the years until retirement eligibility across otherwise similar older workers. Specifically, these changes depend on gender and on small differences in the other two ingredients that determine eligibility, *i.e.* age and years of retirement contributions. As a result, firms with a similar workforce composition experienced differential shifts in the retirement date of employees who were expected to retire soon.

We leverage this feature of the reform to address the endogeneity problem. We restrict our attention to the subset of full-time workers on the cusp of retirement at the time of the reform. We then construct a treatment that measures the average firm-level shift in the full retirement date of these employees. The treatment is only weakly related to the demographics of the firm's workforce in the data; this is because it leverages small idiosyncratic differences across firms in the distribution of age, years of contributions and gender within the narrow subset of employees on the cusp of retirement. Neither does the treatment predict differences in labour demand trends in the pre-reform period. Furthermore, it has a direct economic interpretation as a firm-level shift in the policy parameter, *i.e.* the retirement date of older workers. For this reason, in the second part of the paper we use it to study the short-run effect of raising the full retirement age on the revenues collected from a firm's employees. In the first part of the paper, to document substitutability between workers, we also use the treatment as an instrument to estimate how many individuals are fired or hired when an additional senior employee delays retirement.

To conduct our analysis, we include the treatment into a difference-in-differences model estimated over the period 2009-2015. In the first part of the paper, we compare the labour demand of differentially treated firms before and after the reform. We look at two main margins: layoffs of incumbent workers and external hiring. Our results document that older employees on the cusp of retirement, their co-workers and outsiders are substitutes. Indeed, a larger shift in the full retirement date of older employees leads to more layoffs in the post-reform period. We show that delaying the retirement date of one additional senior worker causes 0.17 more layoffs, 44 per cent of the average number of layoffs in the pre-reform period. Layoffs do not only involve older employees who were expected to retire soon. Young (aged 35 or below), middle-aged (aged 36-55) and other older (aged above 55) co-workers are also fired, causing spillovers within the firm. In particular, the percentage increase in firing is higher for older co-workers than for younger ones, indicating a closer substitutability with older workers expected to retire soon. Delaying the retirement date of an additional older worker also reduces hirings by 0.35 units (7 per cent of the pre-reform average). The decline is largely explained by fewer young and middle-aged new hires under temporary contracts. The effect on dismissals and hiring is concentrated within incumbent workers or outsiders who share the same qualifications (blue-collar, white-collar or manager) as older employees on the cusp of retirement.

In the second part of the paper, we study how the reform affected workers on the cusp of retirement and - due to labour substitutability within the firm - their co-workers. We look at labour earnings and the take-up of social insurance programs. We associate each employee to the firm where he/she worked at the reform date. Then, we aggregate the outcomes of interest across workers of the same type (on the cusp of retirement or co-workers) who were incumbent at the same employer. We find that the pension reform has a positive effect on the earnings of employees close to retirement as they work for longer, but a negative effect on co-workers' earnings. A 1-year shift in the full retirement date of workers close to retirement leads to a 11,356 euros drop in total labour earnings of co-workers in 2015, equivalent to 1.8 per cent of their average value in the pre-reform period. The decline in earnings becomes smaller when we take non-work subsidies into account. Hence, part of the observed dynamics reflects an increased layoff risk. We perform a calibration to infer what share of the drop can be attributed to layoffs. We find that firing explains around one-fifth of the earnings drop. The remaining part of the loss is likely attributable to within-firm earnings dynamics. The latter are more relevant for older and middle-aged employees, consistent with a model where the firm job ladder is based on seniority and the closer substitutes for senior employees on the cusp of retirement are older and middle-aged co-workers.

Stricter criteria for pension eligibility reduce the amount of pension benefits received by employees on the cusp of retirement, but increase - to a lower extent - their take-up of non-work subsidies and disability benefits. The receipt of non-work subsidies also increases among co-workers. Hence, a comprehensive evaluation of the short-run revenues raised by the reform needs to take into account the consequences on workers who are not directly affected by the policy and that arise due to the behaviour of firms. The previous literature has focused on the behavioural responses of older workers (*e.g.* Staubli and Zweimüller 2013; Vestad 2013). We show that spillovers on their co-workers - caused by labour substitutability - are also very important.

To evaluate the implications of labour substitutability for the revenues collected from a firm's employees in the short-run, we develop an accounting model. Our framework allows for spillovers on co-workers and takes into account the effects of the reform on labour tax revenues and other social insurance programs. Using the model, we estimate the fiscal externality of the policy, *i.e.* the share of mechanical savings on the pension payments to a firm's employees that the government loses because of the behavioural responses of these employees and of the firm. According to our estimates, about two-thirds of the savings are lost in the short-run. The cost is entirely explained by spillovers on co-workers. By ignoring these spillovers, we would estimate a positive fiscal externality: higher labour tax revenues on older workers would more than offset costs stemming from their increased take-up of other social insurance programs. We therefore conclude that labour substitutability is pivotal to assessing the consequences of this reform and potentially similar policies that lower the turnover of older workers.

The remainder of the paper proceeds as follows: Section 1.2 reviews the related literature; Section 1.3 illustrates the institutional setting; Section 1.4 describes the data; Section 1.5 outlines the identification strategy; Section 1.6 shows that older workers delay retirement in response to the policy; Section 1.7 presents the main findings on firms' labour demand adjustments; Section 1.8 discusses the main results in relation to the literature; Section 1.9 documents the effect of the reform on co-workers' earnings; Section 1.10 builds a model to estimate the revenues effects of the reform; Section 1.11 concludes.

1.2 Related literature

Our paper relates to the literature that explores the substitutability between workers within the firm using a firm's responses to unforeseen shocks to its workforce. Jäger and Heining (2019) exploit sudden workers' deaths.[4] While they leverage a negative shock to the retention rate, we study a positive one. Unlike a single worker's death, a higher full retirement age may affect the retirement choice of more than one incumbent worker, providing a larger shock to a firm's workforce. Moreover, since our shock affects older employees, it allows us to study substitutabilities between workers within and across age cohorts. We relate our findings to labour demand theory and contribute to the understanding of labour substitutability within firms as studied in models with heterogeneous labour and imperfect labour markets (Cahuc et al., 2008 and Pissarides, 2000). We also add evidence on changes in internal labour market dynamics (Baker et al., 1994).

Our study of the implications of labour substitutability for the revenues raised by

[4]Other recent papers that exploit sudden deaths or hospitalization events to study the importance of directors, CEOs and inventors include Nguyen and Nielsen (2010), Quigley et al. (2017), and Jaravel et al. (2018).

increasing the full retirement age relates to the literature that examines how the generosity of a given social insurance program affects the take-up of other programs.[5] Closely related to our work is Staubli and Zweimüller (2013) who examine a reform increasing the early retirement age.[6] Like them, we show that changes to Social Security rules can generate spillovers on other government programs. We bring a new perspective by treating the firm as an active agent for the transmission of the effects of a pension reform. To do so, we include in our accounting model the demand-driven spillovers on incumbent workers who are not directly affected by the policy in the short-term.

We provide firm-level evidence of substitutability within and across age cohorts. Several papers have studied the relationship between young and older employment within countries or sub-national macro-areas.[7] Gruber and Wise (2010) conclude that the correlation is positive looking at country case studies. More recent work by Bertoni and Brunello (2017) exploits variation in the age structure of Italian provinces and regions. They find that pension reforms causing fewer older workers to retire have negative effects on youth employment, especially during recessionary periods. Exploiting variation in the age structure of the older population across US commuting zones, Mohnen (2019) finds that the retirement slowdown is associated with increased unemployment and occupational downgrading among young individuals. As we discuss later, our results can be regarded as an evidence of the micro-level mechanisms that could deliver substitutability at a macro-level.

Recent and limited literature uses micro-data to investigate how pension reforms that affect elderly labour force participation change labour demand at the firm level. Vigtel (2018) shows that a decrease of the legal retirement age in Norway increases the hiring of senior workers, especially of "risky types" (blue-collar and with previous records of sickness). Martins et al. (2009) study a Portuguese pension reform that increased the legal retirement age for women, finding negative effects on hiring, especially of younger female workers. Boeri et al. (2017) and Carta et al. (2019) evaluate the effects of the *Fornero* pension reform on employment. The former document a reduction of youth employment, whereas the latter find positive effects on young and middle-aged employment in a sample of large firms. Our new contribution is threefold. First, we propose a novel identification strategy that exploits firm idiosyncratic variation in treatment intensity unrelated to broad firm demograph-

[5] Some examples are works on the spillovers of changes in the disability insurance (Karlström et al., 2008; Borghans et al., 2014; Staubli, 2011) or unemployment insurance (Lammers et al., 2013; Inderbitzin et al., 2016). Program substitution effects are also studied by Kline and Walters (2016).

[6] Vestad (2013) also studies how changes to early retirement provisions affect other social insurance programs; Duggan et al. (2007) examine the spillovers on disability insurance of a change in the full retirement age in the US; Atalay and Barrett (2015) explore how increasing the minimum age for women to claim the age pension in Australia affects the take-up of disability benefits.

[7] In a recent strand of literature, Acemoglu and Restrepo (2019) study the interaction of demography and automation showing that robots substitute for middle-aged workers.

ics. Second, we study separately layoffs of incumbent permanent workers and new hires of outsiders; this allows us to explore the susbstitutability among incumbent workers of different ages and the substitutability between incumbent workers close to retirement and individuals outside the firm. Third, we study the relevance of spillovers within the firm to provide novel evidence on the importance of firms for the incidence of these types of policies.

1.3 Institutional Setting

This section describes the main features of the Italian pension system. We provide statistics and institutional details on the Italian labour market in Appendix Section 1.A.

1.3.1 The Italian pension system

As for many OECD countries, including the US, the main pillar of the Italian public pension system is a compulsory pay-as-you-go scheme.[8] A combination of earnings-related and contributions-related methods determines pension benefits. The Social Security tax rate on private payroll employment is 33 per cent. Statutory incidence falls on the employee for one-third and for the remaining two-thirds on the employer.

There are two options for claiming full retirement benefits: old-age pensions and seniority pensions. The eligibility criteria for both options have been substantially redesigned by the 2011 *Fornero* pension reform. The main early retirement option, called *opzione donna*, is available for women only; it allows women to claim benefits before satisfying the old-age or seniority pension requirements. Similarly to early retirement in the US, until 2011 *opzione donna* allowed the claiming of benefits about 3 years before the statutory age.[9] Exercising the early retirement option comes at the cost of receiving lower pension benefits.[10] The average cut is estimated to be roughly 35 per cent of the full benefits that an individual would be entitled to if choosing the seniority retirement option (INPS, 2016). Working past eligibility to retirement is allowed, although protection against dismissals is reduced

[8]In 2015 only 7.3 million people (less than one-third of people employed) had supplementary private pension plans (COVIP, 2015). Although the number has been growing in recent years, the limited diffusion of these financial instruments is partly due to the fact that the public pension system is relatively generous.

[9]Early retirement using *opzione donna* was possible in 2011 upon turning 57 years old (with 35 years of contributions). In the same year, private sector female employees could claim an old-age pension upon turning 60 years old (with 20 years of contributions).

[10]This is in part due to the fact that pension benefits are calculated using a fully contributions-based method, rather than a combination of the contributions-based and the, typically more generous, earnings-based method.

when employees reach a certain combination of age and years of contributions.[11]

1.3.2 Statutory and observed retirement date

How retirement choices change in response to pension reforms that extend the full retirement date determines the extent to which firms experience an increase in the retention rate of older employees when these policies are implemented. We define an individual as retired when he/she starts collecting retirement benefits. Retirement spikes around the earliest among statutory full retirement dates in our data; more than 70 percent of individuals retired within a 1-year window around such a date in 2012 (Appendix Figure 1.H.1). This trend is common to other countries. In the US the share of workers retiring at the full retirement age has been increasing in the last decade; the share of individuals exercising the early retirement option has also starkly dropped (Munnell and Chen, 2015). Estimates in Mastrobuoni (2009) for the US document a strong response of retirement choices to the full retirement date: an increase in the full retirement age by 2 months increases the age at retirement by around half as much. A similar response emerges from our data where a one year shift in the full retirement date translates into a 7-month delay in retirement (see Section 1.6).

1.3.3 The *Fornero* reform

The *Fornero* pension reform was passed in December 2011 (Decree Law 201/11, article 24). It was part of the “Save Italy” decree, an emergency package of measures in response to the mounting pressure of financial markets on the Italian sovereign debt. Designed by a new technocratic government and approved one month after its appointment, it entered into force in January 2012. Although the need for a deficit reduction package was anticipated, its exact content was not known in advance. Moreover, the decision and implementation lags were both very short. As a result, anticipatory effects were likely negligible. The reform raised the requirements for claiming old-age and seniority pensions, reducing the number of new retirees and increasing the average age at retirement.[12] The new dispositions applied to all workers who did not accrue the right to claim retirement benefits under pre-reform rules by the end of 2011. A few other categories of workers - listed in Appendix Section 1.B - were grandfathered. For all other private sector employees, Table 1.1 compares the main features of pre- and post-reform retirement rules over the period

[11] In particular, the age and years of contributions are those needed by men for claiming an old-age pension.

[12] Appendix Figures 1.H.2 plots the evolution of age at retirement and of the number of new retirees over the period 2005-2015, distinct by gender and type of pension (old-age and seniority). It focuses on retirees who were private sector workers. The participation rate in the 55-64 age group increased from 39.3 per cent in 2011 to 57.0 per cent in 2018, halving the gap with respect to the average in euro-area countries.

2012-2015, for old-age (Panel a) and seniority (Panel b) pensions. Appendix Table 1.H.1 and Appendix Section 1.B provide further details.

Old-age pensions: The reform raised the age requirement for old-age pensions, whilst leaving the contribution requirement (20 years) unchanged. In 2011 the statutory retirement age was 60 for women and 65 for men. Absent the reform, it would have risen gradually to reach 61 years and 10 months for women and 65 years and 7 months for men in 2018. Per effect of the reform, the age requirement has further increased, to reach 66 years and 7 months for both genders in 2018. Hence, the change in the age requirement was considerably larger for women than for men.

Seniority pensions: The reform re-designed the rules for claiming seniority pensions. Until 2011 a "quota" system was in place; workers could retire as soon as their age and years of contributions summed up to a "quota", conditional on both surpassing a certain threshold. In 2011 the quota was set to 96, conditional on being at least 60 years old and having at least 35 years of contributions. Had the rules not changed, the "quota" would have gradually risen from 96 to 97.6 over the 2012-2108 period.[13] Alternatively, workers could retire upon totalling 40 years of contributions, regardless of their age. The *Fornero* reform abolished the "quota" system. It legislated that in 2012 a seniority pension could be claimed upon totalling at least 41 years and 1 month of contributions for women and 42 years and 1 month of contributions for men; the requirement gradually increased to reach 41 years and 10 months and 42 years and 10 months in 2018, respectively.[14]

The reform did not abolish the early retirement option. The take-up of such option was very low until 2011, because of the cut in benefits. After the reform, which substantially tightened eligibility criteria for claiming full benefits, the take-up of *opzione donna* increased. Yet, even in the year when it peaked within our sample period (2015), less than 20 per cent of eligible women claimed early retirement. Moreover, only 80 per cent of them made job-to-retirement transitions (INPS, 2016).[15] As a result, the take-up of *opzione donna* remains limited in our sample.

The reform caused heterogeneous changes in years until retirement eligibility among otherwise similar older workers. As a result, firms with a similar senior workforce were affected by the reform to a different extent. Appendix Figure 1.H.3 shows the relationship between age and years of contributions in 2011 and the shift in the

[13]The age requirement for the "quota" 97.6 would have been 61 years and 7 months.

[14]If individuals start collecting retirement benefits before age 62, there is a 1 per cent penalty for every year between the ages of 60 and 62, which raises to 2 per cent for every year before the age of 60.

[15]The remaining 20 per cent were unemployed or out of the labour force when they retired.

retirement date.[16] It focuses on workers who could have started collecting pension benefits by 2014 under pre-reform rules. The heatmap highlights the substantial variation in the extent of the shift. Among female workers, the most affected are those between 58 and 59 years old with less than 35 years of contributions in 2011. Their retirement age shifts by 4 years or more.[17] Women with more than 36 years of contributions or who are closer to 60 years old experience smaller changes. Among male workers, those close to eligibility under "quota 96" are the most affected. Milder changes apply to male employees who would have retired under the old-age option or who were close to reaching 40 years of contributions.

1.4 Data

We leverage high-quality and confidential administrative data available at the Italian Social Security Institute (*Istituto Nazionale della Previdenza Sociale*, INPS). We describe below the three main sources that we combine to build the dataset used for the analysis.

Workers' contribution histories: we have access to previously unexploited contribution histories for all employees (more than 6 million) who worked in small-medium sized firms between 2009 and 2015.[18] For every contribution spell in any given year we observe the following information: (i) the number of weeks covered by contributions to Social Security; (ii) the event triggering the payment of contributions (*e.g.* paid work, maternity leave, unemployment benefits) and the amount of contributions paid. We use the information on the length of contribution spells to compute the number of accrued years of contributions by the end of 2011, which is needed to predict the change in the full retirement date. The information on the type and amount of contributions allows us to derive measures of the take-up of non-work subsidies as well as comprehensive measures of earnings, which include labour income from quasi-salaried employment, self-employment and public sector jobs.

Matched employer-employee records: we also exploit linked employer-employee records available over the period 1983-2015 for the universe of non-agricultural firms with at least one employee. Firms report detailed information about employees cov-

[16]In Subsection 1.5.1.1 we describe how we compute the expected retirement dates under pre- and post-reform rules.

[17]Before the reform, they were close to retiring under the "quota 96" option. After the reform, their earliest available retirement option became either the new seniority pension, which requires more years of contributions, or the old-age pension, the age requirement of which had been substantially raised by the reform.

[18]We focus on firms with 3 to 200 employees in the first quarter of 2009. The restriction stems from limitations to the maximum number of workers' contribution histories that could be made available by the INPS for the sake of the project.

ered by Social Security filling the so-called *Uniemens* modules. The data covers 74 per cent of private employment and 93 per cent of private sector employees.[19] We use monthly data for the period 2009-2015.[20] Firms are identified by a unique Tax Identification Number (TIN). In case of a multi-establishment firm, all establishments feature the same TIN. For a subset of firms the TIN is associated to multiple Contribution Identification Numbers, which however do not necessarily identify the different establishments. In our analysis we will focus on firms with a single Contribution Identification Number.

For each worker-firm record, the following information is available: (i) beginning and end date of the contract, alongside the underlying motivation (*e.g.* layoff, quit); (ii) type of contract (permanent vs fixed-term, full-time vs part-time); (iii) broad occupation group (blue-collar, white-collar or manager); (iv) wage; (v) number of days worked.[21] We link these records to workers' and firms' registers containing baseline information, such as gender and age of employees as well as opening date, sector and location of businesses. Drawing on these sources, we build yearly firm-level measures of adjustments in labour demand by counting total new hires and layoffs of permanent workers. We also construct these measures for different categories of workers, as identified by their contract, occupation or demographic group.

Register of retirees: the register of retirees provides information about the type of pension paid to each retiree, including disability benefits, as well as the date when the first pension payment was collected and the amount received. We use this information to study how the reform affected the receipt of old-age and seniority pensions, as well as of disability benefits.

1.5 Empirical Strategy

The purpose of our empirical analysis is twofold. First, we aim to study the firm-level substitutability between workers on the cusp of retirement, their co-workers and outsiders, by exploiting a reform that delayed the retirement of older employees. Second, we want to evaluate the implications of substitutability for the short-run revenues that the government raises on a firm's employees through this policy. Addressing these questions is challenging because the share of workers retained in a firm due to the reform is strongly related to the firm's demographics. Variation along

[19]Self-employment accounts for most of the share of total private employment that the data does not cover. The agricultural sector accounts for most of the missing share of private sector employees.

[20]The INPS has been collecting matched employer-employee records with an annual frequency since 1983 and with a monthly frequency since 2005. Since our analysis spans the period 2009-2015, we mostly use the latter dataset, relying on the former to compute worker-level measures of experience and tenure.

[21]We deal with multiple records for worker-firm-month combinations using the procedure explained in Appendix Section 1.C.

this dimension could reflect differences in unobservable labour demand dynamics. Hence, it cannot be directly exploited for identification. We instead leverage the fact that the reform changed the full retirement date of similar older workers to a different extent. Specifically, our treatment exploits variation in the firm-level average shift in the full retirement date of employees on the cusp of retirement. This variable changes across firms due to idiosyncratic differences in the distribution of gender, age and years of contributions among the narrow set of these individuals, but does not reflect substantial variation in the broad demographics of the firm. It also captures shifts in the policy parameter (the full retirement age), the revenues effect of which we examine in the second part of the paper. In the first part of the analysis, we provide a more intuitive measure of substitutability that shows how an increase in the number of older employees who delay retirement affects the number of workers hired or fired. To do so, we use our treatment in an instrumental variable setting.

1.5.1 Treatment construction

To construct the treatment, we follow a two-step procedure. First, we compute the change in the full retirement date for employees close to retirement before the reform, who we define *potential retirees* (Subsection 1.5.1.1). Second, we construct the average variation in the full retirement date of *potential retirees* employed at a given firm when the reform is passed (Subsection 1.5.1.2). We exploit such an identifying variation within a difference-in-differences regression framework (Subsection 1.5.2).

1.5.1.1 Individual shift in the full retirement date

For the purpose of computing the retirement date, an older worker can be summarized by his/her type $\theta(g, a, c)$, where g is gender, while a and c are age and years of contributions as of December 2011, respectively. We draw on workers' demographics to build the first two variables and on contribution histories to compute total years of contributions, following the rules detailed in Appendix Section 1.D. For every type θ we compute the reform-induced change in years until full retirement, thus excluding early retirement options. To this end, we construct the predicted retirement dates according to pre- and post-reform rules and denote the difference with δ_θ:[22]

$$\delta_\theta = \text{Years until full retirement date}_\theta^{\text{post-reform}} - \text{Years until full retirement date}_\theta^{\text{pre-reform}} \tag{1.1}$$

If early retirement choices are influenced by the reform, δ_θ is an individual assignment to treatment as opposed to the actual change in the retirement date. To construct δ_θ, we take as given the contribution history observed in the data up to 2011 and we make the following assumptions on the post-2011 contribution history:

(i) workers accrue full contributions on their accounts (52 weeks per year) until retirement;

(ii) the predicted retirement date is the earliest date at which the worker can start collecting benefits under either the old-age or the seniority pension scheme.

Assumption (i) requires that individuals work year-round and full-time in the post-reform period.[23] Data shows that the median annual contribution is 52 weeks for workers aged 60 or above in 2012, suggesting that assumption (i) has solid ground. Assumption (ii) provides a criterion to select among the different options for claiming full benefits: after predicting the retirement date associated with every available option, we select the earliest one.[24] As discussed earlier, the literature has documented that retirement behaviour displays bunching at the acquisition of full pension rights and Appendix Figure 1.H.1 shows consistent evidence for Italy.

For all workers expected to retire by 2014 under pre-reform rules and who actually retired by 2017, we compare actual and predicted retirement dates. We define the actual retirement date as the date when an individual starts collecting Social Security benefits. Appendix Figure 1.H.4 shows a "forecast quality" assessment. The majority of the differences between the two dates (69 percent) lies within a 1-year window, indicating that our measure is quite accurate in predicting the actual retirement date (panel a). This also provides supportive evidence to assumption (ii), because a thin right tail of the distribution implies that workers do not often retire later than we predict. The left-skewness arises because of two main reasons. First, women can use the early retirement option (*opzione donna*), which causes a larger difference between actual and predicted retirement dates for female workers (panel

[22] Before the *Fornero* reform abolished them, the so-called "waiting windows" (*finestre mobili*) were in place: workers could start collecting pension benefits only 12 months (or more, see Appendix Table 1.H.1) after becoming eligible. Most workers postponed retirement until that moment. We incorporate this feature of pre-reform rules by assuming that, had the reform not passed, employees would have retired when they could start collecting benefits.

[23] Alternatively, we require that non-work periods are covered by *figurative* contributions (see Appendix Section 1.D).

[24] Under this assumption, the option to claim full benefits may change because of the reform: some types θ could first reach eligibility to an old-age pension under pre-reform rules, while the seniority pension may become the earliest option available under post-reform rules, and viceversa.

b). Second, some workers maintain the right to retire under pre-reform rules (see Appendix Section 1.B).

1.5.1.2 Firm-level shift in the full retirement date

We focus on the change in the full retirement date of older workers on the cusp of retirement before the reform. We define *potential retirees* as the full-time employees who could have retired within 3 years under pre-reform rules (*i.e.* by 2014) and who are directly affected by the reform in the short-run. The 3-year threshold also allows us to focus on a subset of workers with similar age and contribution histories, who at the same time face a diverse enough variation in their full retirement date because of the reform. We will also show how results change when considering shifts in the full retirement date of workers further away from retirement in 2011. Panel (a) of Figure 1.1 plots the distribution of the change in years until the full retirement date for *potential retirees* employed in the sample of firms that we describe in Subsection 1.5.2. As already emerged from the heat-maps in Appendix Figure 1.H.3, δ_θ displays a substantial variability, with the mean equal to 1.36 and the standard deviation equal to 1.4.

Every *potential retiree* of type θ experiences the same shift δ_θ in the expected full retirement date. To construct the firm-level change in the full retirement date, we build a shift-share shock. We weight the δ_θs by the share of every θ in the subset of *potential retirees* employed in the firm. The firm-level treatment T_i therefore is:

$$T_i = \sum_{\theta \in \text{Potential retirees}} \pi_{\theta,i}\, \delta_\theta \tag{1.2}$$

$\pi_{\theta,i}$ is the share of type-θ workers among *potential retirees* employed at firm i in the last quarter of 2011. The $\pi_{\theta,i}$s depend neither on firm size nor on the share of *potential retirees* out of the firm's total workforce. As a consequence, we show that they do not reflect meaningful differences in the broad demographics of the firm. T_i therefore captures the idiosyncratic firm-level shift in the full retirement date of *potential retirees*. It also has a straightforward interpretation as the change in the policy parameter that was shifted by the reform. Because of these properties, we use T_i as the treatment variable. In the first part of the paper, where we aim to study the extent of substitutability among workers, we also use T_i as an instrument for the number of older workers who delay retirement.

The distribution of T_i for firms that employ at least one *potential retiree* displays significant variability. The mean is 1.37 and the standard deviation is 1.33 (Figure 1.1, panel b). The variability of T_i decreases with firm size.[25] This is one of the

[25]This is because larger firms employ more *potential retirees*, so that the distribution of age, gender and contributions among them is more likely to mirror the one prevailing in the universe of employees on the cusp of retirement. For the same reason, T_i does not exhibit substantial variation

reasons why we focus on small and medium firms, where the variability of T_i is greater.

1.5.2 Empirical Specification and identifying assumptions

1.5.2.1 Empirical specification

To study the effect of the reform we restrict our attention to the sample of firms with a single Contribution Identification Number and with 3-200 employees in the first quarter of 2009 that are active in every year between 2009 and 2015.[26] We then focus on firms that employ at least one *potential retiree* in the quarter when the reform is passed (last quarter of 2011). Firms with no *potential retirees* may not be an appropriate control group, because they have a different demographic composition and are likely to differ along other unobserved time-varying characteristics. We nonetheless show that results are qualitatively similar in the universe of firms in the 3-200 size class. These restrictions leave us with a balanced panel of 61,434 firms.

In this sample, we estimate a difference-in-differences model with a continuous treatment and multiple pre- and post-reform periods:

$$Y_{it} = \lambda_i + \gamma_t + \sum_{k=2009}^{2015} \beta_k^T \mathbb{1}(k = t) \cdot T_i + \varepsilon_{it} \tag{1.3}$$

i indexes the firm and t indexes the year. Y_{it} is the outcome of interest. λ_i is a firm fixed effect that captures time-invariant heterogeneity across firms, including differences in average outcomes across treatment levels; γ_t are year fixed effects that control for year-specific shocks common to all firms. Standard errors are clustered at the firm level to address the potential concern of serial correlation across periods (Bertrand et al., 2004). The coefficients of interest are $\{\beta_k^T\}_{k=2009}^{k=2015}$ and show how the treatment T_i affects the outcome of interest in year k relative to the reform year.[27]

To summarize the results, we also estimate a more compact difference-in-differences regression comparing pre-reform years (2009-2011) to post reform years (2012-2015):

$$Y_{it} = \lambda_i + \gamma_t + \beta^T Post_t \cdot T_i + \varepsilon_{it} \tag{1.4}$$

We interact the treatment with the dummy $Post_t$ that takes value 1 in years 2012-2015. β^T captures the average effect of a 1-year increase of T_i in the post-reform years.

To provide a measure of the extent of labour substitutability, in the first part of

across local labour markets.

[26] The majority of firms in our size range have a single Contribution Identification Number; some firms have multiple Contribution Identification Numbers, which however do not necessarily map into different establishments.

[27] We set β_{2011}^T equal to 0.

the analysis we also re-scale the coefficients so that they capture the effect of an additional *potential retiree* delaying retirement. Specifically, we estimate a version of specification (1.4) where we use T_i as an instrument for the number R_i of *potential retirees* who delay retirement by at least one year relative to the predicted full retirement date under pre-reform rules.[28] R_i therefore is:

$$R_i = \sum_{j:j \in \text{Potential Retirees}_i} \mathbb{1}(\tilde{\delta}_j \geq 1) \tag{1.5}$$

where $\tilde{\delta}_j$ is the observed change in the full retirement date of individual j, *i.e.* the difference in years between the observed retirement date and the predicted pre-reform full retirement date.

1.5.2.2 Identification assumptions

We leverage variation in the characteristics of *potential retirees* for identification. The extent to which firms are affected by the reform depends on the distribution of the shares of types θ among their *potential retirees* (equation 1.2). Identification requires that $\pi_{\theta,i}$s do not correlate with firms' unobservable time-varying characteristics (Borusyak et al., 2018 and Goldsmith-Pinkham et al., 2019). In other words, the characteristics of *potential retirees* employed at a given firm should not correlate with a firm's time-varying unobservables. The shares $\pi_{\theta,i}$s depend neither on firm size nor on the share of *potential retirees*. Thus, we leverage a source of variation that does not depend explicitly on the firm's broad demographics; we only exploit the variability that stems from idiosyncratic differences across firms in the distribution of gender, age and years of contributions (*i.e.* in the distribution of types θ) among the narrow set of *potential retirees*. Evidence that the composition of *potential retirees* in a firm relates to trends in labour demand would provide a sign of potential threats to identification. Pre-trends as captured by the coefficients $\{\beta_k^T\}_{k=2009}^{k=2011}$ provide suggestive evidence of the conditional exogeneity of T_i. If trends are parallel among differentially treated firms, these coefficients should not be different from zero. We also perform placebo and balancing tests to assess the validity of the identifying assumptions.

1.5.2.3 Placebo tests

We assess whether T_i predicts labour demand trends in the pre-reform period by running a series of placebo tests. To this end, we artificially assign the date in which the reform becomes effective to 2010 or 2011, rather than to 2012. We then estimate a version of specification (1.4) on the pre-reform period (2009-2011). We

[28]We exploit the fact that, everything else being equal, larger shifts in the full retirement date result in more *potential retirees* substantially delaying retirement. The exclusion restriction requires that the change in the full retirement date affects firms only through a delay in the retirement of *potential retirees*.

test the effect of T_i on layoffs and new hires, which are the main firm-level outcomes we study in Section 1.7. Appendix Table 1.H.2 shows that a one standard deviation (hereafter 1σ) increase of T_i has virtually zero effects. It is only marginally significant in predicting layoffs when the placebo reform is assigned to 2009. This indicates that firms facing heterogeneous shifts in the full retirement date of *potential retirees* in 2011 did not exhibit substantially different demand trends in the pre-reform period. Appendix Table 1.H.2 also shows that the standardized share of *potential retirees* delaying retirement out of a firm's employment (*i.e.* R_i/N_i) predicts trends in labour demand in a stronger and statistically significant way in three out of four cases. This provides ground to the choice of not exploiting variation that captures differences in the broad demographic composition of firms.

1.5.2.4 Balancing tests

We run balancing tests whereby we regress a rich set of firm characteristics at the beginning of the period (first quarter of 2009) on T_i, while controlling for province-sector fixed effects. Appendix Table 1.H.3 shows that the correlations are very weak, although precisely estimated. This holds true when looking at the gender and age composition of the firm. Despite the fact that the reform affected on average women more than men, a 1σ increase of T_i is associated with a decline in the share of male employees of only around 1 percentage point, against an average of 66 per cent. The coefficients relative to the share of young, middle-aged and older workers are 0.00, -0.01 and 0.01, against averages of 0.30, 0.58 and 0.12. We nonetheless show in Subsection 1.7.6 that the results on labour demand adjustments do not change when including the interactions between firm baseline characteristics and year fixed effects as controls. Moreover, column 2 of Appendix Table 1.H.3 shows that the correlations between firm baseline characteristics and R_i/N_i are stronger in almost all cases, further suggesting that firms with heterogeneous demographics differ along many other dimensions.[29]

1.5.3 Descriptive Statistics

Appendix Table 1.H.4 compares the characteristics of firms in our main sample to other firms with a single Contribution Identification Number, in the same size class and active in 2009-2015. Firms with at least one *potential retiree* are on average three times as large as other firms and older. They are more likely to operate in the manufacturing sector and have a higher share of blue-collar workers. Their employees are older, more experienced, have higher tenure and higher daily wages. Appendix Table 1.H.5 shows that *potential retirees* are older, more experienced, and have higher tenure than other full-time employees working in our sample of firms. They have higher gross daily wages, and are more likely to have a permanent

[29]For comparability T_i and R_i/N_i are standardized.

contract. Appendix Table 1.H.6 compares *potential retirees* to other older employees who are similar along many dimensions.[30] Before the reform, employees closer to retirement were 5 percentage points more likely to be absent from work because of sickness and 1 percentage point more likely to be absent due to work-related injuries or sick leave, suggesting a decline in effort as workers approach retirement.

1.6 Older workers delay retirement after the reform

The reform may affect a firm's labour demand if it prolongs the working lives of older workers. We first investigate the response of retirement choices to changes in the full retirement date by estimating an individual-level version of specification (1.3) on the sample of *potential retirees* employed in our sample of firms. We use the worker-level shift in the full retirement date as the treatment (δ_θ in equation 1.1).[31] Figure 1.2 shows that delaying the full retirement date by one year causes an increase in time spent at work and a decline in time spent on retirement.[32] A 1-year increase in δ_θ is associated with 1.5 more months worked and 2.3 less months in retirement in 2015. The increase in months spent at work is smaller than the decrease in months spent on retirement, suggesting that some workers are no longer employed when they start claiming pension benefits. The decline in time spent on retirement is smaller than it would have been had if all workers retired at the predicted post-reform full retirement date ("benchmark"). The difference between the two lines reflects early retirement responses to the shift in the full retirement date.

To quantify the response of retirement choices to the policy, we also use the same treatment δ_θ in a cross-section regression where the outcome is the difference between the observed retirement date and the pre-reform predicted full retirement date (Table 1.2). The specification includes age and gender fixed effects. The coefficient on δ_θ captures by how much workers delay retirement when they face a 1-year shift in the full retirement date. Individuals delay retirement by 7.1 months in response to a 1-year delay in the full retirement date. The response is slightly larger for men (7.51 months against 6.97 for women). These figures are close to estimates in Mastrobuoni (2009) for the US, where a 1-year increase in the normal retirement age causes a 6-month increase in the age at retirement.

[30]We perform a coarsened exact matching procedure. Matching covariates are: age, gender, type of contract (open-ended vs fixed-term), occupation, as well as firm's province, sector and size.

[31]The regression features individual fixed effects. We also add a vector of controls including gender and age dummies interacted with year fixed effects.

[32]Appendix Table 1.H.7 summarizes the results estimating an individual-level version of specification (1.4) on the sample of *potential retirees* employed in our sample of firms.

1.7 Labour demand responses to the reform

In this section we document how labour demand for co-workers and outsiders changes when there is a shift in the full retirement date of *potential retirees* employed at the firm. We first present estimates based on specification (1.3); we then summarize the effects by estimating a single post-reform coefficient based on the more compact specification (1.4). Coefficients show the effect of a 1-year increase in T_i. Finally, we present estimates from a modified version of specification (1.4), where T_i is used as an instrument for the number R_i of *potential retirees* who delay retirement by at least one year. This allows us to provide a more direct measure of the extent of substitutability between employees on the cusp of retirement, co-workers and outsiders.

According to labour demand theory, a drop in demand for other workers caused by the retention of older employees can be reconciled with complementarity between the two types of labour only in the case of an increase in other workers' wages. We provide the intuition in Appendix Section 1.E, where we develop a labour demand model with heterogenous labour and we incorporate different wage formation mechanisms. As we document below, a shift in the full retirement date of *potential retirees* reduces the demand for co-workers and outsiders. A large wage increase is inconsistent with the evidence in Subsection 1.9.1 that shows a negative effect on the earnings of incumbent co-workers. These results suggest that incumbent co-workers are substitutes for older workers close to retiring. Since the fall in labour demand also involves younger co-workers, our findings provide evidence of substitutability across workers of different ages within the firm. In Section 1.8 we discuss how our results relate to the vast literature on aggregate employment patterns of different age cohorts.

1.7.1 Layoffs

Firms increase layoffs of permanent employees in response to a shift in the full retirement date of *potential retirees*.[33] Panel (a) of Figure 1.3 shows results from specification (1.3).[34] Pre-reform coefficients are not significantly different from 0, providing supporting evidence to the assumption of parallel trends until 2011. In the post-reform period all coefficients are positive: the effect of an increase in T_i on layoffs is small and not significant in 2012, when most *potential retirees* would have worked even under pre-reform rules; it then grows and becomes significant in 2013 and 2014, when most *potential retirees* would have retired in absence of the reform. In 2015, when some of the *potential retirees* eventually retire, the growth

[33]Since labour regulations force firms to pay a temporary worker until the contract end date if he/she is fired for economic reasons, the cheapest way to part from a temporary employee is not to renew their contract. Thus, we focus on layoffs of permanent workers.

[34]Appendix Table 1.H.8 shows the estimates reported in all panels of Figure 1.3.

in the estimated effect flattens out. In that year a 1-year increase in T_i leads to 0.055 more layoffs - 14 per cent of the pre-reform average number of layoffs per year (0.39). Table 1.3 reports results from the more compact specification (1.4). In the first row it shows that a 1-year increase in T_i causes on average 0.037 more layoffs per year in the post-reform period (column 1); a 1σ increase in T_i (a shift of 1.33 years in the full retirement date) is therefore associated to 0.05 more layoffs, which amounts to 12.6 per cent of the pre-reform average. In the second row we re-scale the coefficients by using T_i as an instrument for R_i. The first stage coefficient is positive (0.21), statistically significant, and the associated Kleibergen-Paap rk Wald F-statistics is large. For every additional *potential retiree* who delays retirement by at least one year, firms fire on average 0.17 more workers per year in the post-reform period, which amounts to 43.6 per cent of the pre-reform average number of layoffs.

To study the substitutability between incumbent workers, Panel (b) of Figure 1.3 breaks down the effect by workers age. We classify as young the employees of age 35 or below, as middle-aged those aged between 36 and 55, and as old those aged above 55. Layoffs increase in all age groups due to the reform, providing the first evidence that firings do not only involve *potential retirees*. Comparing the coefficients in columns 2 to 4 of Table 1.3 to the group-specific average number of layoffs before the reform, the percentage increase is highest among older workers and lowest among young employees. Focusing on re-scaled coefficients, an additional *potential retiree* who delays retirement by at least one year causes 0.086 more layoffs of middle-aged employees and 0.04 more layoffs of old employees; these amount to 43 and 66.7 per cent of the pre-reform averages, respectively.

We further investigate whether the layoffs adjustment on older workers is concentrated within *potential retirees* or also involves their co-workers (Panel (c) of Figure 1.3 and column 5 of Table 1.3). When focusing only on older workers who are not *potential retirees* the coefficients halve, suggesting that half of the response is concentrated on this category of workers. The firm's median share of *potential retirees* out of older workers is 66 per cent. Thus, *potential retirees* are not disproportionately affected relative to other older employees. Since *potential retirees* and other older workers earn similar wages, the cost of firing the two types of workers is the same.[35] We find that firms trade them off in a similar fashion, suggesting a strong substitutability.

Panel (d) of Figure 1.3 shows estimates when the dependent variable is the ratio between the number of layoffs in every age group and the respective number of incumbent employees at the beginning of the period (first quarter of 2009).[36] The

[35]The cost of layoffs - when ruled unfair by a labour court - is a function of the fired worker's wage.

[36]We therefore estimate this specification on the subset of firms that have at least one employee in each age group in that quarter-year.

coefficients therefore can be interpreted as an approximation of the effect of T_i on the probability of being laid-off. Older workers face a larger probability of being fired than younger co-workers in the post-reform period.

1.7.2 New Hires

Firms more affected by the reform reduce hiring in the post-reform period relative to less affected firms, while there is no evidence that they were on different trends before the reform (Panel (a) of Figure 1.4).[37] The pattern of post-reform coefficients is u-shaped: the negative effect emerges as early as in 2012 and peaks in 2013, when a 1σ increase of T_i leads to 0.14 fewer new hires, which amounts to a little less than 3 per cent of the pre-reform average (4.79). In 2015 the coefficient is still negative, but not significantly different from 0. This suggests that firms delay hiring in response to the reform; new hires decline in its aftermath and bounce back as *potential retirees* become eligible to retire under post-reform rules. Table 1.4 shows that a 1σ increase of T_i on average reduces hiring by almost 0.1 per year (column 1). Re-scaling the coefficients, we find that for every additional *potential retiree* who delays retirement by at least one year, hiring declines on average by 0.35 unit per year, which amounts to 7.3 per cent of the pre-reform average.

Panel (b) of Figure 1.4 decomposes the effect by new hires' age. The decline in hiring affects young and middle-aged outsiders to a similar extent. On the other hand, there is no significant effect on older workers, who constituted a small share of new hires in the pre-reform period. For every additional *potential retiree* who delays retirement by at least one year, young and middle-aged new hires decline on average by 0.20 and 0.15 units per year, respectively (columns 2 to 4 of Table 1.4). The decline only concerns new hires employed with temporary contracts (Panel (c) of Figure 1.4 and columns 5-6 of Table 1.4). In 2015 the close to 0 effect on total hiring masks substantial heterogeneity; the effect on new hires under fixed-term contracts is still negative, while it becomes positive for permanent contracts.[38]

1.7.3 Do firms respond to retirement delays happening in the future?

All workers experience an increase in years left to full retirement, except those who become eligible by the end of 2011.[39] We have studied the effects of shifting the full retirement date of workers on the cusp of retirement. In this subsection we test whether firms respond to changes in the retirement date of workers who were less close to retire at the time of the reform. Specifically, we estimate our specifications on the sample of firms that employ no *potential retirees*, but have at

[37] Appendix Table 1.H.9 shows the estimates reported in all panels of Figure 1.4.

[38] Part of this could be caused by a generous package of incentives to foster the use of permanent contracts that was available to firms in 2015.

[39] See Section 1.3.3 and Appendix Section 1.B for a list of the other few grandfathered groups.

least one worker expected to retire in 2015 or 2016 under pre-reform rules. The treatment is defined as the average shift in the full retirement date of these workers. Appendix Figure 1.H.5 shows that layoffs do not respond in a significant way. There is a negative effect on new hires starting from 2013, but it is only significant in one year. Appendix Table 1.H.10 confirms that there is a virtually null effect on layoffs and a negative, but not significant, effect on new hires. Overall, there is no clear evidence that in the immediate aftermath of the reform firms respond to retirement delays that involve workers who were not close to retiring.

1.7.4 Labour demand adjustments within and across occupation groups

We further explore how the shock to the retirement date of *potential retirees* is absorbed within the firm by looking at the decisions of its units. We call unit the group of employees in a specific occupation (blue-collar, white-collar or manager). We estimate versions of (1.3) and (1.4) at the firm-unit level. Our specifications separately include a treatment defined within the group of *potential retirees* in the unit as well as one defined within the group of *potential retirees* in other units of the same firm. The within-unit treatment generates a larger effect on layoffs and new hires (Appendix Figure 1.H.6 and Appendix Table 1.H.11). Thus, a shift in the full retirement date of *potential retirees* employed in a given unit impacts co-workers and outsiders in the same occupation group; on the other hand, spillovers across units in the same firm are limited. Indeed, Appendix Table 1.H.11 shows that for both outcomes we can reject the hypothesis that the two treatment effects are equal. Our evidence is consistent with higher substitutability between workers who perform similar tasks. Similarly, Jäger and Heining (2019) find that labour substitutability is higher within occupation groups.

1.7.5 Heterogeneity by pre-reform turnover rates

Firms with a higher propensity to separate from workers should be more prone to adjust labour demand in response to shocks. To test this, we construct a measure of a firm's turnover rate by using the average number of separations in the pre-reform period divided by employment at the beginning of the period (first quarter of 2009). We define as separations all events of layoffs, terminations of temporary contracts and voluntary quits. We then split firms into two groups (high and low turnover) based on whether they fall above or below the median of the distribution of the turnover measure. We add a triple interaction to our specification where a high turnover dummy is interacted with the treatment and year fixed effects.[40] The effect on layoffs is significantly larger in high-turnover firms and the difference relative to low-turnover firms is statistically significant (Appendix Figure 1.H.7 Panel (a) and

[40]We also interact the high turnover dummy with the set of year fixed effects, to estimate a fully interacted model.

Appendix Table 1.H.12). Thus, high-turnover firms manipulate more easily the margin of layoffs. We also look at new hires and we find some evidence that the negative effect of the reform is larger in high-turnover firms (Panel b). However, the difference in the treatment effect between the two groups of firms is not statistically significant (Appendix Table 1.H.12).

1.7.6 Sensitivity checks

Appendix Figure 1.H.8 shows that results are robust to a battery of robustness checks. We start by verifying that our estimates do not capture differential labour demand trends across firms which have different characteristics at the beginning of the period. We showed in Appendix Table 1.H.3 that the relationship between T_i and firm baseline characteristics is weak. Nonetheless, we address this concern by augmenting specification (1.3) with the interaction between year dummies and a vector of firm covariates measured in the first quarter of 2009. First, we include in the vector of covariates a set of dummies for the quintiles of the share of female employees in the firm, because the pension reform affected women to a greater extent than men. Adding these controls reduces the concern that non-parallel labour demand trends across firms with a different gender composition confound our estimates. Second, we add to the vector of covariates dummies for quintiles of several other variables: firm size, firm age, average wage, the share of young (≤ 35), middle-aged ($36-55$) and older (>55) workers. Third, we estimate a specification where year fixed effects are interacted with two-digit sector and province fixed effects to check that our estimates are not confounded by heterogeneous economic cycles across sectors and provinces. Estimates based on these augmented specifications are very similar to the main ones. Fourth, we add one year (2008) to the pre-reform period to show that labour demand trends were similar up to four years before the reform in a longer balanced panel. Finally, we check whether results are robust in the universe of firms employing between 3 to 200 employees in the first quarter of 2009; for this purpose, we set $T_i = 0$ if a firm employs no *potential retirees*.[41] This exercise reveals that, while heterogeneously treated firms appear to be on parallel trends in terms of firings before the reform, this is less the case for hiring, consistent with our conjecture that firms employing a younger workforce are not a good counterfactual for firms employing *potential retirees*. Nonetheless, the post-reform coefficients are similar - if anything larger - to those coming from the baseline specification.

We also check whether results on layoffs could be confounded by other policy changes. The 2012 *Fornero* labour market reform and the 2015 *Jobs Act* have gradually reduced the costs of unfair dismissals for firms with more than 15 employ-

[41] Also in this case, we restrict the attention to firms that remain active between 2009 and 2015 and have a single Contribution Identification Number.

ees.[42] We check whether firms below and above this threshold behave differently. Appendix Figure 1.H.9 shows that the percentage increase in layoffs in firms with at most 15 employees at the beginning of the period (first quarter of 2009) is very similar to that of firms with 16 to 30 employees. We conclude that changes in the cost of layoffs do not seem to be a relevant confounder for our analysis.

1.8 Discussion

In this section we discuss general equilibrium effects and the relationship between our results and the previous literature.

General equilibrium effects: We conducted a partial equilibrium analysis of the short-run firm responses to the reform. However, general equilibrium dynamics could affect our identification threatening the implicit assumption on the absence of spillovers across firms. The responses of labour demand and supply to a higher retirement age may affect market tightness and the outside option of different cohorts of workers. Hence, firms that are not directly affected by the reform can change their behaviour because of spillovers caused by other firms. However, it takes time for these dynamics to come into effect, so that general equilibrium effects are not likely to play a significant role over the horizon of our analysis (*i.e.* the first 4 years after the implementation of the reform). The fact that new hires and layoffs respond in the immediate aftermath of the reform, when general equilibrium effects are likely less relevant, further reduces these concerns. In addition, reform-induced reductions in the number of new hires are significant, but tiny when compared to total new hires and cyclical fluctuations in hiring.[43] Finally, we showed in Subsection 1.7.3 that new hires in the aftermath of the reform respond mostly to the shift in the full retirement date of *potential retirees*, while future retirement delays do not seem to have an effect. This reduces the extent to which firms' responses could affect the labour market general equilibrium in the four years after the reform.

Relation to the literature on age substitutability: The *Fornero* reform provides a neat experiment to document how employers trade-off workers of different ages within the firm. We exploit a firm-level idiosyncratic shock to uncover firms' labour complementarities, and our results show evidence of substitutability among age cohorts within firms. This firm-level substitutability contributes to equilibrium allocations in local labour markets, but it is not sufficient to draw definitive

[42] Appendix Section 1.A provides details about these two policies, discussing their implications for the cost of dismissals.

[43] An additional *potential retiree* who delays retirement by at least one year causes a 0.35 drop in the number of new hires. There are 61,434 firms in our sample and the average value of R_i is 0.83. A back of the envelope calculation delivers a drop in total new hires of 17,850 units per year, 0.2 per cent of total new hires in 2011 and less than 5 per cent of the average yearly fluctuation in hires in the period 2012-2015.

conclusions about the relationship between the employment dynamics of different cohorts as studied by Gruber and Wise (2010). Documenting age complementarities in labour markets requires economy-wide demographic shocks instead of firm-level ones, and a long-run perspective that allows for slow adjustments in workers' outside options and employment flows. For instance, a recent work by Mohnen (2019) looking at decade-long horizons shows that the retirement slowdown of older workers is associated with worse labour market outcomes for younger workers in US counties. Similar results have been found by Bertoni and Brunello (2017) for Italian provinces and regions. Our results can be regarded as a first step in uncovering the micro-level mechanisms that explain patterns of macro-level substitutability among age cohorts.

1.9 Workers' earnings and take-up of other social insurance programs

In this section we study how the reform affects *potential retirees* and - due to labour susbtitutability - their co-workers. The information contained in the contribution histories allows us to track workers across jobs and to observe their take-up of various social insurance programs. We can therefore study the consequences of the policy for all employees who are incumbent in our sample of firms at the reform date, including workers who leave the firm in the post-reform period. We focus our attention on full-time employees and we document that the reform led to an increase in labour earnings of *potential retirees*, whereas it caused a decline in labour earnings of their co-workers. We then examine the take-up of disability and unemployment insurance. The literature has documented that older workers substitute away from pension benefits into other social insurance programs in response to reforms that tighten the criteria for pension eligibility (*e.g.* Duggan et al., 2007, Staubli and Zweimüller, 2013 and Atalay and Barrett, 2015). We find similar evidence. Spillovers on younger co-workers caused by labour substitutability within the firm have received less attention in the literature. This section documents these spillovers as a preliminary step to the quantification of the short-run revenues collected by the reform from a firm's employees.

1.9.1 Earnings of Incumbent Employees

We start by focusing on co-workers earnings. We define co-workers as all full-time employees who were not expected to retire within 3 years when the reform is passed and who worked in a firm with at least one *potential retiree*. We match every co-worker to the firm where he/she was incumbent at the quarter of the reform. We then estimate specification (1.3) on our sample of firms using the total of co-workers' labour earnings as the dependent variable; specifically, Y_{it} is the total labour earnings in year t of all co-workers who were incumbent in firm i in the last quarter of 2011.[44] Labour earnings include income from other private-sector employers, self-employment, and public-sector employment.[45] The reform leads to a decline in incumbent co-workers' earnings, which grows over time (Figure 1.5).[46] A 1-year shift in the full retirement date of *potential retirees* at the firm causes a drop of 11,356 euros in 2015, which is equal to 1.8 per cent of average total labour earnings pre-reform. When adding non-work subsidies to labour earnings, the decline becomes smaller (1.3 per cent in 2015). Table 1.5 summarizes the results based on specification (1.4), reporting an average yearly drop equal to 6,550.7 euros, reducing to 4,902 when non-work subsidies are taken into account (columns 1 and 2).

Columns 3 to 5 of Table 1.5 show results by co-workers age, focusing on those who had permanent contracts. It emerges that middle-aged and older co-workers are the most affected. The average drop in earnings amounts to 4,567 and 1,521 euros per year respectively (1 per cent and 3.6 per cent of the pre-reform average labour earnings); the loss experienced by young co-workers amounts to 1,220 euros per year (0.9 percent). In line with the findings on layoffs described in the first part of the analysis, adjustments seem to be borne more by older co-workers than younger ones.

Since the negative effect on co-workers earnings moderates after accounting for non-work subsidies, the increase in layoffs documented in Subsection 1.7.1 can partly explain this. Within-firm dynamics could also play a role if incumbent co-workers experience earnings reductions or a slower earnings growth when *potential retirees* delay retirement. To quantify the relative contribution of these two channels, we perform a decomposition exercise. We combine estimates of the effects of the reform on layoffs with estimates of the cost of a job loss. The ideal experiment to estimate the effect of being laid-off would randomize such an event across workers. In absence of this experiment, we match every worker fired after the reform to workers who do not separate from their firm in the same period.[47] We perform a coarsened exact match (CEM) along several covariates. To assess the cost of layoffs, we then estimate a difference-in-differences regression on the matched sample whereby the treatment is a dummy equal to 1 if the worker is fired.[48] The estimated earnings drop is 5,455

[44]We therefore collapse observations from the worker-year level to the firm-year level, by aggregating workers who were incumbents at the same firm in the quarter when the reform was passed. Regressions are weighted based on firm size at the beginning of the period (first quarter of 2009).

[45]Earnings are winsorized at the 99th percentile.

[46]Appendix Table 1.H.13 reports all the coefficients displayed in the figures discussed in this subsection.

[47]We only consider the first layoff event that causes a separation from the firm where the worker was incumbent in the last quarter of 2011.

[48]We add to the specification the matching covariates interacted with time fixed effects. The covariates are age, sex, wage, occupation, type of contract, experience, sector, province, firm size. We also weight controls based on the standard CEM weights (see Iacus et al., 2012). We discuss the weighting and further details about the match in Appendix Section 1.F.

euros three years after the separation (Appendix Figure 1.H.10). It amounts to a 19.1 per cent decline, in line with estimates in Couch and Placzek (2010).[49]

Appendix Figure 1.H.11 presents the result of our calibration exercise. The blue-shaded area is the share of the total earnings loss of full-time permanent workers that we can impute to layoffs. It accounts for around 17 per cent of the total effect of the reform on earnings, suggesting that although layoffs play a relevant role, within-firm dynamics provide a major contribution. We replicate the exercise for young, middle-aged, and older workers (Appendix Figure 1.H.12). Some important heterogeneity emerges. Separations explain a larger part of the earnings losses for young workers; 25 per cent of the total as opposed to 17 per cent for middle-aged and 11 per cent for older workers. This heterogeneity is consistent with a model of seniority where earnings grow with age within the firm, creating career spillovers across workers (Bianchi et al., 2019). As a consequence, older co-workers are the closest substitutes to *potential retirees* and bear the largest adjustments in term of earnings dynamics.

We also study earnings dynamics for *potential retirees*. As we did for co-workers, we match every *potential retiree* to the firm where she was incumbent at the reform quarter and we sum up labour earnings across all *potential retirees* incumbent in the same firm. Figure 1.5 and Table 1.5 column 6 show an increase in their labour earnings, which is the natural consequence of a prolonged working life.

1.9.2 Pension benefits and spillovers to other social insurance programs

We also study whether shifts in the full retirement date affect the take-up of pension benefits as well as of other social insurance programs. Specifically, we focus on non-work subsidies and disability benefits. We aggregate individual-level outcomes at the firm level following the procedure described in Subsection 1.9.1. While the previous literature on program substitution has mostly studied the responses of workers directly targeted by a given policy change (in this setting, *potential retirees*), we also examine the response of co-workers, who can be affected due to labour substitutability within the firm.

Figure 1.6 reports estimates based on specification (1.3).[50] Focusing on *potential retirees* (Panel a), the reform-induced shift in the full retirement date significantly lowers the receipt of pension benefits. As T_i increases by 1 year, the amount of pension entitlements received by all *potential retirees* who were incumbent in firm i drops by 1670 euros in 2012; the decline grows over time to reach 7,000 euros in

[49]Couch and Placzek (2010) revisit pioneering work by Jacobson et al. (1993). They find that the earnings loss for displaced workers is around 30 per cent after one year and 9 per cent six years after the dismissal. See also Davis and von Wachter (2011), Farber (2017), and Schmieder et al. (2018) for more recent estimates of the cost of job loss.

[50]Table 1.H.14 displays estimates based on the more compact specification (1.4).

2015 and the cumulative drop in the first four years after the implementation of the reform amounts to almost 19,000 euros.[51] On the other hand, transfers from other social insurance programs increase: the cumulative raise in the 2012-2015 period is equal to 1,253 euros, much smaller in absolute value than the decline of pension benefits, and reflects higher enrollments both in the disability insurance program and - to a larger extent - in the unemployment insurance program. These findings are similar to those of Staubli and Zweimüller (2013), who study the consequences of an increase in the early retirement age in Austria.

With regards to co-workers (Panel b), the amount of pension benefits received in the first four years after the passing of the reform changes by a negligible and non significant amount, which is consistent with the fact that these workers were not eligible to retire soon under pre-reform rules. On the other hand, delaying the retirement of *potential retirees* leads to an increase of social insurance transfers to co-workers. The cumulative increment in the 2012-2015 period associated to a 1-year increase in T_i is equal to roughly 7,000 euros, larger than that documented for *potential retirees*. It reflects almost entirely the increased receipt of non-work subsidies, while the effect on the take-up of disability benefits is small.

1.10 Implications of substitutability for the revenues raised on incumbents

Our results document that - due to labour substitutability - the reform caused large spillovers on all incumbent workers. As a way to prove their relevance in the analysis of policy incidence, we develop a model to estimate the implications of these spillovers for the revenues collected on incumbent full-time employees in the short-run. We estimate the share of public pension savings raised mechanically on incumbents, which is lost due to the behavioural responses of a firm and of its workers. This is the fiscal externality that stems from the channels that we observe. We incorporate into our model the spillovers caused by labour substitutability, which have so far received little attention in the literature. We then discuss the contribution of our results to the analysis of the welfare effect of the reform and their relevance for other public policies that reduce the incentives of older workers to leave firms.

1.10.1 An accounting model

We construct a model of government accounting that considers two types of agents defined as in our empirical analysis with the labels of *potential retirees* (p)

[51]The cumulative change is computed as the sum of post-reform coefficients estimated using specification (1.3) (*i.e.* $\sum_{k=2012}^{2015} \beta_k^T$). Cumulative changes are reported in Appendix Table 1.H.15.

and co-workers (c). Agents perform different labour-related activities. The main activity is paid labour in a firm. A positive share of workers receives transfers in the form of non-work subsidies, disability benefits, or pension entitlements. The budget constraint for individual i is:

$$x_i \leq (1 - \tau_i)\, w_i\, l_i^w + \hat{T}_i \tag{1.6}$$

x_i is consumption. Labor l_i in a firm is paid a wage w_i and taxed at rate τ_i. The worker receives total transfers $\hat{T}_i$ that depend on time spent in different labour and non-labour activities, including retirement. We describe the details of all the components of $\hat{T}_i$ in Appendix Section 1.G.

The fiscal externality is the share of mechanical revenues raised through the reform from a firm's employees that is lost because of behavioural responses:

$$FE = -\frac{\text{Cost of Behavioral Responses}}{\text{Mechanical Public Pension Savings}} \tag{1.7}$$

The formula is derived in Appendix Section 1.G. The numerator represents the costs incurred by the government because of behavioural responses. These costs occur when extra non-work subsidies are paid to fired employees, more workers enroll into disability insurance, and lower tax revenues are raised from labour income when incumbent workers experience a drop in earnings. Mechanical revenues in the denominator are the resources that the government would save from a firm's employees because of the policy, absent any change in the behaviour of the workers and the firm. Hence, they measure the savings that would arise if every *potential retiree* retired at the post-reform full retirement date. When FE is between -1 and 0 the reform generates an increase in the revenues collected. If the fiscal externality falls below -1, the government loses the entire mechanical revenues on incumbents because of behavioural responses. This is the case of local Laffer effects (Hendren, 2019 and Werning, 2007).

Our framework is highly stylized and ignores some of the general equilibrium effects of the policy. The model abstracts from the revenues lost on marginal workers who are not hired due to the reform. However, we provide estimates of these losses based on conservative assumptions in our calibrations. Due to the lack of balance-sheet information, we cannot incorporate the effect of the reform on a firm's performance. To the extent to which the reform affects revenues and profits, our model misses their externalities on the government budget. We also lack information on *potential retirees* and other workers who are not employed in a firm at the time of the reform, but on whom the reform generates mechanical savings in pension outlays.[52] The

[52]If a delayed full retirement age increases labour force participation and some of these workers find a job, extra revenues could be raised by taxing their labour earnings (Carta et al., 2019).

reform could also affect workers' health creating externalities for the public health system, for which we do not have data. Finally, our analysis focuses only on small and medium firms as in the rest of the paper and on full-time workers.

1.10.2 Empirical implementation and results

The fiscal externality depends on how the reform affects labour earnings and social insurance transfers (Section 1.9). We provide a detailed description of the empirical implementation in Appendix Section 1.G, with a discussion on the alternative calibrations that we implement.

Table 1.6 presents the main results. Standard errors of the estimates are bootstrapped via a block bootstrap procedure with 1000 repetitions.[53] We start by computing the fiscal externality following the standard approach in the literature. This exercise illustrates what would happen if we ignored the spillovers on co-workers that are caused by labour substitutability and focused only on the restricted sample of *potential retirees* who are directly affected by the reform. The first column of Table 1.6 shows that all estimates are positive, indicating that the government raises more money than what is mechanically saved through the policy on a firm's incumbent workers. Since the *potential retirees* work for longer per effect of the reform, they increase labour earnings (Figure 1.6). Tax revenues from the extra earnings are sufficient to offset increased transfers from unemployment and disability insurance programs.

We add the spillovers on co-workers to the model in columns 2 to 4 and we show the fiscal externality estimates across alternative calibrations of the average tax rate and of the early retirement pension benefits. Point estimates range from -0.56 to -0.66, indicating that, even when we add spillovers, the savings on pension outlays overcome the cost of behavioural responses. However, labour demand spillovers generate - and entirely explain - a non negligible loss of mechanical revenues. The reason is twofold. First, co-workers are affected because of labour substitutability, suggesting its importance for assessing the impact of this reform. In particular, they experience an increase in non-work subsidies as a consequence of layoffs. Moreover, the reform has a negative effect on their labour earnings, which results in lower labour tax revenues. Second, mechanical savings in pension outlays are raised only on the workers who were expected to retire within our horizon of analysis. As Figure 1.6 suggests, savings only come from *potential retirees*, who represent a small share of the workers in the sample. Over a longer time horizon, a larger share of the workforce will contribute to generating mechanical savings, increasing the revenues from

[53] We perform a block bootstrap that corrects residuals using the wild bootstrap procedure introduced by Wu et al. (1986), Liu et al. (1988) and Mammen et al. (1993). This procedure allows us to obtain asymptotic refinement for standard errors when residuals are correlated within firm and *iid* across firms.

the policy.

We conclude our analysis by extending the model to provide a more conservative estimate of the fiscal externality. We have so far disregarded the tax revenue losses on marginally non-hired workers. To provide an estimate of these losses, we assume that every marginally non-hired worker would earn no labour income for as long as the median duration (3 months) of a job search for individuals who find a job in 2012-2015. We calibrate earnings losses using the median labour earnings of new hires in the 3-month period following the hiring event and we calibrate the number of marginally non-hired workers using estimates from the first part of the analysis.[54] In this conservative scenario, the fiscal externality is larger (column 5). It ranges from -0.65 to -0.70, indicating that about 2/3 of the mechanical savings are lost in the short-run.

Our results show that spillovers within the firm significantly affect the revenues raised on incumbent full-time employees. Hence, the behaviour of firms and labour substitutability are crucial to studying the incidence of this reform. Substitutability between age cohorts has potentially important implications for other policies that affect the incentives of older workers in a similar way. Examples are an increase in the early retirement age, higher penalties for early retirement, higher monetary incentives for working after the full retirement date, changes in the eligibility criteria for disability insurance. All these policies extend the time that older employees spend at work, increasing their retention at the employer firm. The response of firms to an increased retention will affect those incumbent workers who are substitutes for older workers, creating unintended spillovers qualitatively similar to those documented in our analysis.

1.10.3 Substitutability, spillovers and welfare in the short-run

We conclude by discussing how firms' behaviour and labour substitutability may influence the welfare effects of the policy in the short-run. The reform welfare impact has two components: the first is the fiscal cost of the policy, the second is the workers marginal willingness to pay for the policy. In the long-run, the reform will generate large pension savings, and the workers marginal willingness to pay will largely depend on the extent to which future cohorts benefit from an extension of the retirement age.[55] We cannot estimate these long-run components because we lack data on an extended time horizon. Even if the data was available, conclusions would strongly depend on general equilibrium effects that are hard to disentangle.

Yet, we can highlight how the short-term components of welfare are affected by

[54]See Appendix Section 1.G for more details about this calibration.

[55]Even though some workers could be willing to pay to avoid the increase in the full retirement date, the willingness to pay can be high for some others since the policy improves the sustainability of the Social Security system.

labour substitutability and firms' decisions. First, the workers marginal willingness to pay for the reform is affected by the spillovers caused by adjustments in labour demand; firm responses have first-order utility effects, which depend on how much workers value employment. Estimates of the latter are hard to obtain, but we expect these effects to be relevant given the increase in co-workers' layoffs. Second, we quantify part of the other welfare component, by measuring the short-run revenues collected by the reform on incumbent employees. By estimating the fiscal externality, we show that spillovers - caused by substitutability - have sizable effects on public finances.

1.11 Conclusions

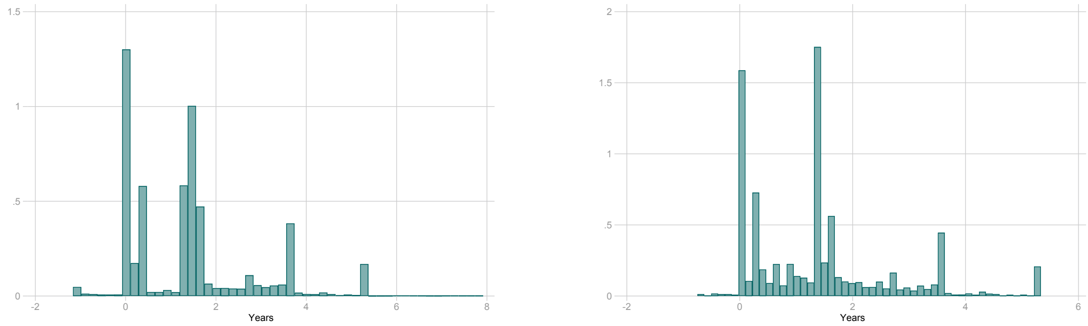
This paper studies the importance of labor substitutability and firm's decisions when investigating the effects of an increase in the full retirement date. In contexts where the response of retirement choices to the full retirement age is high, the most proximate consequence of this policy for a firm is the increase in the retention of workers on the cusp of retirement. We develop a novel empirical strategy to measure the reform-induced shock that is particularly effective for small and medium firms. We show that labor demand responds to the reform in a way that is consistent with substitutability between workers on the cusp of retirement, their co-workers, and outsiders. Older co-workers are the closest substitutes for senior employees who delay retirement. Spillovers within the firm also have significant implications for the revenues raised on incumbent employees through the reform. They cause all of the revenue losses in the first four years after implementation, indicating that labor substitutability and firm's decisions play an important role in studying the welfare effects of this and of similar policies.

Since the reform was implemented during a budget crisis, as with many similar policies, the fiscal externality that we estimate is a first step to quantifying the effects of the policy on the short-run government budget. Even though firm's responses might have been amplified by adverse economic conditions, the contribution of spillovers to the revenues raised on incumbents is unlikely to vanish in more favorable phases of the business cycle. In light of these findings, we argue that firms are an important vector for the passthrough of Social Security policies and thus should be considered in welfare calculations. Clearly, our estimates cannot be directly extrapolated to other contexts. Yet, our results on substitutability may extend to other policies that lower the incentives of older workers to leave a firm. Examples are increases in the early retirement age (*e.g.* Staubli and Zweimüller, 2013) or stricter eligibility criteria for disability insurance (*e.g.* Staubli, 2011).

1.12 Figures

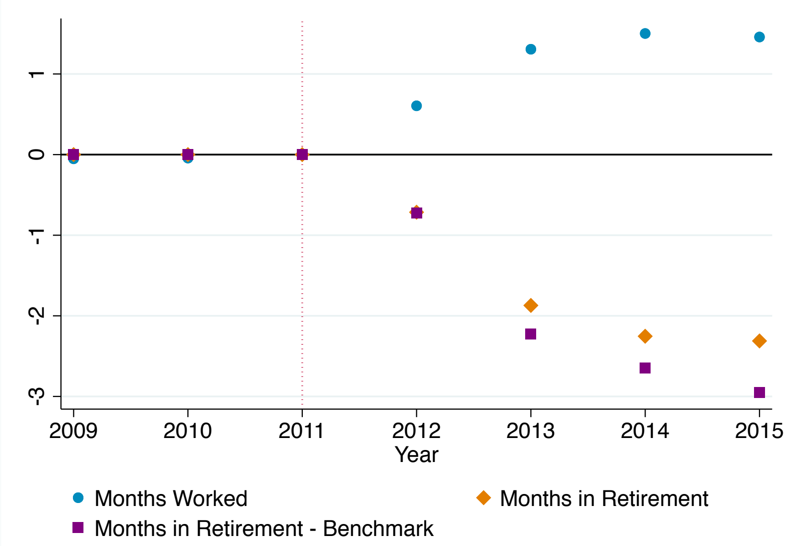
Figure 1.1: The distribution of the shift in the full retirement date

(a) Worker-level (b) Firm-level



Notes: Panel (a) plots the distribution of the individual-level shift in the full retirement date of *potential retirees* in our sample of firms (Subsection 1.5.3). Panel (b) shows the distribution of T_i, the firm-level average shift in the full retirement date of *potential retirees*, excluding the first and last percentiles. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. The predicted full retirement dates of *potential retirees* under post-reform rules are capped at December 2020, as dispositions published in 2012 did not span a longer horizon. The capping, nonetheless, only applies to very few individuals. Due to the abolition of the "waiting window" few workers face a negative change, i.e. they can retire sooner under post-reform rules. Number of workers = 98,358. Worker-level shift mean = 1.36 (std. dev. = 1.4). Number of firms = 61,434. Firm-level shift mean = 1.37 (std. dev. = 1.33).

Figure 1.2: The effect of the reform on the working life of *potential retirees*



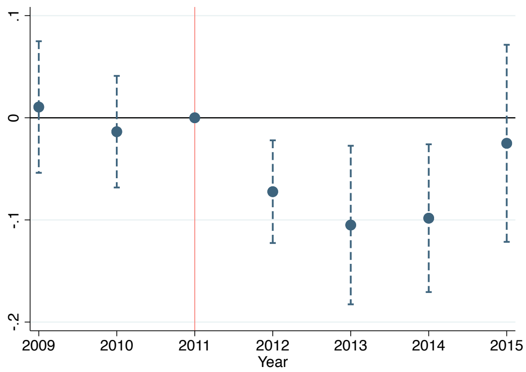
Notes: The figure plots estimates from an individual-level version of specification (1.3), where the unit of analysis is the single *potential retiree* in our sample of firms, the treatment is δ_θ and there are age-year and gender-year fixed effects. Standard errors are clustered at the individual level. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Dots, diamonds and squares represent the effects of a 1-year increase in the full retirement date on months worked, observed months in retirement and predicted (benchmark) months in retirement if all workers retired at the predicted post-reform full retirement date. We define an individual as retired when he/she starts collecting retirement benefits. Whiskers represent 95 per cent confidence intervals, not available for months spent in retirement due to the null variability of this outcome before 2011. Number of observations = 688,492. Pre-reform mean outcomes: months worked = 11.28, months in retirement = 0.00, months in retirement (benchmark) = 0.00.

Figure 1.3: Labour demand adjustments: layoffs

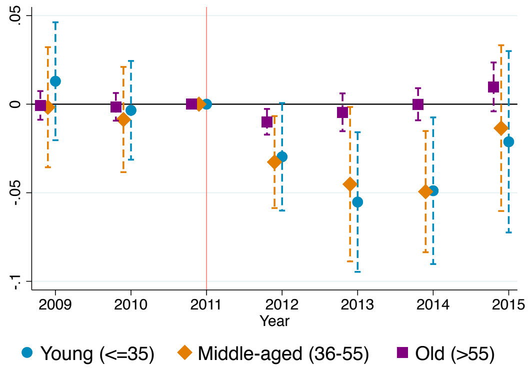
Notes: The figure reports the effect of a 1-year increase of T_i on total layoffs (Panel a), layoffs by age group (Panel b), layoffs of older workers (Panel c) and layoffs probability by age group (Panel d). Whiskers represent 95 per cent confidence intervals. Estimates are based on specification (1.3), which includes firm and year fixed effects. Standard errors are clustered at the firm level. The layoffs probability in Panel (d) is defined as the ratio between the number of layoffs in every age group in a given year and the respective number of incumbent employees at the beginning of the period. T_i is the average change in the full retirement date of *potential retirees* employed at a given firm when the pension reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Young workers are aged 35 or below, middle-aged workers are between 36 and 55 years old, old workers are over 55 years old and old non-*potential retirees* are old workers who in 2011 were not expected to retire within 3 years under pre-reform rules. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. In Panel (d) the sample of firms is further restricted to include only firms with at least one worker in each age group at the beginning of the period. Number of observations = 430,038 (305,319 in Panel d). Mean outcome pre-reform: total = 0.39; young = 0.13; middle-aged = 0.2; old = 0.06, old non-*potential retirees* = 0.04.

Figure 1.4: Labour demand adjustments: new hires

(a) New hires



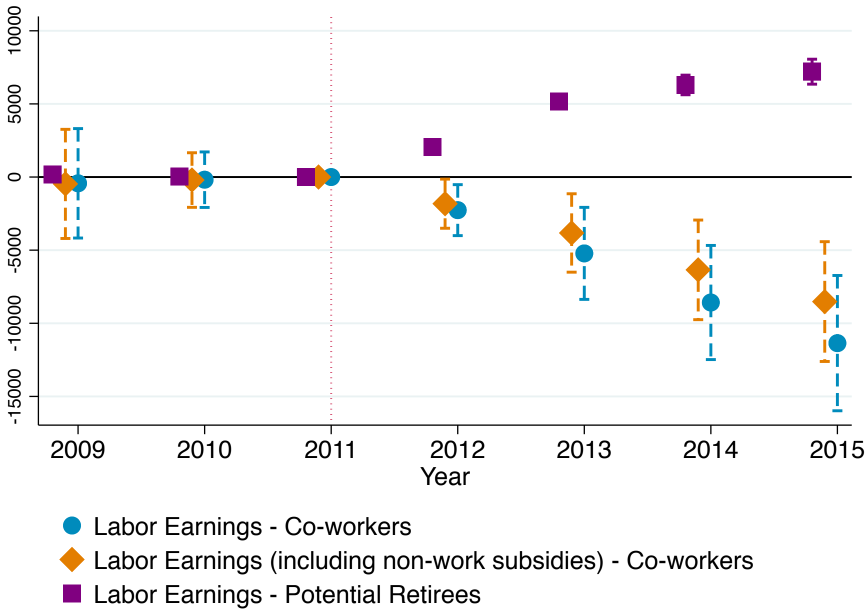
(b) New hires by age



(c) New hires by type of contract

Notes: The figure reports the effect of a 1-year increase of T_i on total new hires (Panel a), new hires by age group (Panel b) and new hires by type of contract (Panel c). Whiskers represent 95 per cent confidence intervals. Estimates are based on specification (1.3), which includes firm and year fixed effects. Standard errors are clustered at the firm level. T_i is the average change in the full retirement date of *potential retirees* employed at a given firm when the pension reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Young workers are aged 35 or below, middle-aged workers are between 36 and 55 years old, old workers are over 55 years old. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. Number of observations = 430,038. Mean outcome pre-reform: total = 4.79, young = 2.38, middle-aged = 2.06, old = 0.35, permanent = 1.46, temporary = 3.32.

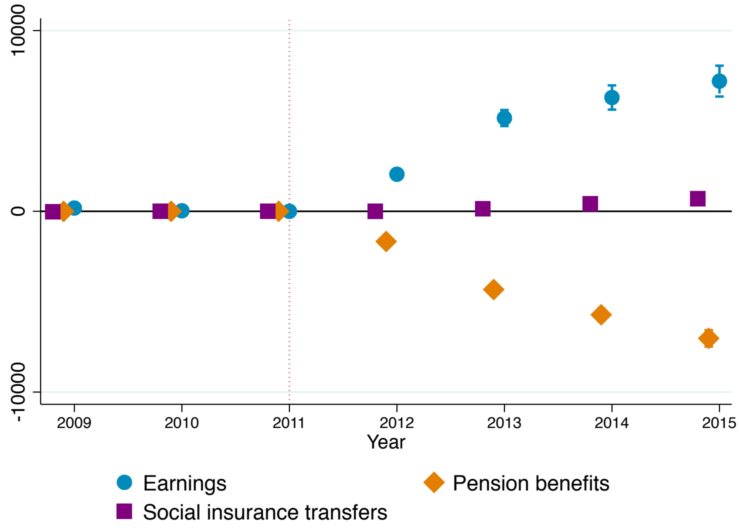
Figure 1.5: Incumbents' labour earnings



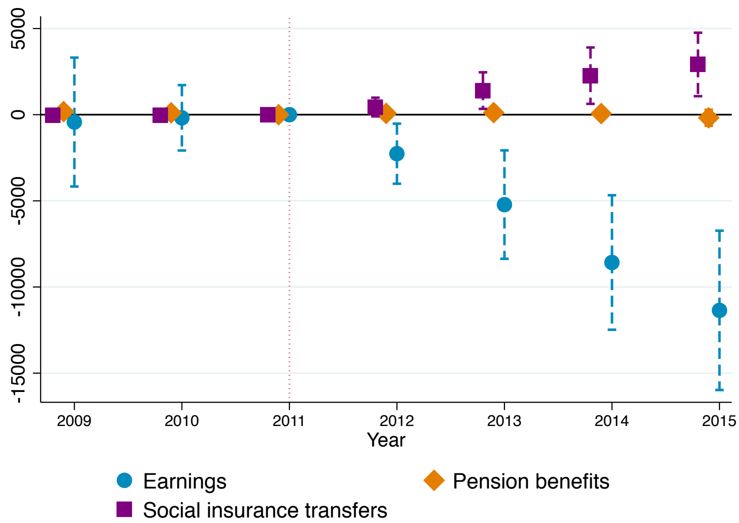
Notes: The figure shows the response of labour earnings of *potential retirees* and of their full-time co-workers to a 1-year increase in T_i, the shift in the full retirement date of *potential retirees* employed at the firm when the reform is passed. Whiskers represent 95 per cent confidence intervals. Standard errors are clustered at the firm level. Worker-level observations are aggregated at the level of the firm where workers were incumbents in the last quarter of 2011. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Labour earnings are winsorized at the 99th percentile; non-work subsidies are winsorized at the 99.99th percentile. The regressions are based on specification (1.3) and are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. Observations are weighted according to firm size at the beginning of the period. Number of observations = 423,346 (425,971 in the regression for *potential retirees*). Mean labour earnings pre-reform: co-workers = 647,005.96; *potential retirees* = 49,951.32.

Figure 1.6: Incumbents' labour earnings, pension benefits and social insurance transfers

(a) *Potential retirees*



(b) Co-workers



Notes: The figure shows the response of labour earnings, pension benefits and total social insurance transfers (non-work subsidies and disability benefits) to a 1-year increase in T_i, the shift in the full retirement date of *potential retirees* employed at the firm when the reform is passed. Whiskers represent 95 per cent confidence intervals. Standard errors are clustered at the firm level. Labor earnings, pension benefits and disability benefits are winsorized at the 99th percentile; non-work subsidies are winsorized at the 99.99th percentile. Worker-level observations are aggregated at the level of the firm where workers were incumbents in the last quarter of 2011. See notes to Figure 1.5 for further details on these regressions. Number of observations = 423,346 (425,971 in the regression for *potential retirees*) Mean outcomes pre-reform: *potential retirees'* labour earnings = 49,951.32; *potential retirees'* pension entitlements = 4.2; *potential retirees'* total transfers = 303.21; co-workers' labour earnings = 647,005.96; co-workers' pension entitlement = 1,337.52; co-workers' total transfers = 1,702.16.

1.13 Tables

Table 1.1: Pre- and post-reform pension requirements

	Panel (a): Old-age pension			
	Men		Women	
	Pre-reform	Post-reform	Pre-reform	Post-reform
2011	65YA	Not in place	60YA	Not in place
2012	65YA	66YA	60YA	62YA
2013	65YA+3MA	66YA+3MA	60YA+3MA	62YA+3MA
2014	65YA+3MA	66YA+3MA	60YA+4MA	63YA+9MA
2015	65YA+3MA	66YA+3MA	60YA+6MA	63YA+9MA

	Panel (b): Seniority pension		
	Pre-reform	Post-reform	
		Men	Women
2011	Quota 96 (60YA and 35 YC) or 40 YC	Not in place	
2012	Quota 96 (60YA and 35 YC) or 40 YC	42YC+1MC	41YC+1MC
2013	Quota 97.3 (61YA+3MA and 35 YC) or 40 YC	42YC+5MC	41YC+5MC
2014	Quota 97.3 (61YA+3MA and 35 YC) or 40 YC	42YC+6MC	41YC+6MC
2015	Quota 97.3 (61YA+3MA and 35 YC) or 40 YC	42YC+6MC	41YC+6MC

Notes: The table shows the pre- and post-reform requirements for old-age and seniority pensions, over the period 2012-2015. YA and MA flag the age requirement in terms of years and months, respectively. YC and MC flag the contribution requirement in terms of years and months, respectively. Additional details can be found in Appendix Table 1.H.1.

Table 1.2: Response of retirement choices to the change in the full retirement date

	All (1)	Men (2)	Women (3)
δ_θ (years)	7.07*** (0.05)	7.51*** (0.08)	6.97*** (0.06)
Observations	98,355	69,368	28,977
Mean δ_θ	1.36	1.27	1.58
Std. Dev. δ_θ	1.4	1.07	1.95

Notes: The table reports estimates from a cross-section regression where the outcome is the difference (in months) between the observed retirement date and the expected retirement date under pre-reform rules. The regression also includes age and gender fixed effects. Standard errors in parentheses are clustered at the province×sector level. The treatment is the individual-level change in years left to retirement caused by the reform (δ_θ). The sample consists of all *potential retirees* in our sample of firms. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. We define an individual as retired when he/she starts collecting retirement benefits. If a worker has not retired by the end of 2017 (the last month of available data on pension claims), we assume that he/she will retire in January 2018. Column (1) shows the results for all *potential retirees*, column (2) and (3) show the results for male and female *potential retirees*, respectively.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.3: Labor demand adjustments: layoffs

	All	Young	Middle-aged	Old	Old *Non-potential retiree*	Layoffs probability Young	Layoffs probability Middle-aged	Layoffs probability Old
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$T_i \cdot$ Post	0.037***	0.010***	0.018***	0.009***	0.004***	0.004***	0.003***	0.005***
	(0.006)	(0.002)	(0.004)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$R_i \cdot$ Post	0.174***	0.047***	0.086***	0.040***	0.021***	0.020***	0.015***	0.023***
	(0.030)	(0.009)	(0.019)	(0.006)	(0.005)	(0.006)	(0.006)	(0.004)
Mean outcome pre-2012	0.39	0.13	0.2	0.06	0.04	0.15	0.24	0.07
Observations	430,038	430,038	430,038	430,038	430,038	305,319	305,319	305, 319
Mean T_i	1.37	1.37	1.37	1.37	1.37	1.44	1.44	1.44
Std. Dev. T_i	1.33	1.33	1.33	1.33	1.33	1.38	1.38	1.38
Mean R_i	0.83	0.83	0.83	0.83	0.83	0.93	0.93	0.93
Std. Dev. R_i	0.99	0.99	0.99	0.99	0.99	1.08	1.08	1.08
1st stage coeff.	0.21***	0.21***	0.21***	0.21***	0.21***	0.20***	0.20***	0.20***
1st stage F-stat.	8,288.71	8,288.71	8,288.71	8,288.71	8,288.71	5,072.32	5,072.32	5072.32

Notes: The table reports estimates from specification (1.4), which includes firm and year fixed effects. Standard errors in parentheses are clustered at the firm level. Column 1 shows the effect on layoffs of all workers, columns 2 to 4 on young ($\leq$ 35 years old), middle-aged (aged 36-55) and old (above 55 years old) workers, respectively. Column 5 shows the effect on old employees who are not-*potential retirees*, while columns 6 to 8 report the effect on the probability of layoff by age. The layoffs probability is defined as the ratio between the number of layoffs in every age group in a given year and the respective number of incumbent employees at the beginning of the period. In the first row the treatment is T_i, the average change in the full retirement date of *potential retirees* employed at the firm when the pension reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. In the second row the treatment is R_i, the number of *potential retirees* who retire at least one year after the predicted pre-reform full retirement date, and T_i is used as an instrument. The first stage F statistics in the IV regression is the Kleibergen-Paap rk Wald F statistics. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. In columns 6 to 8 the sample is further restricted to include only firms with at least one worker in each age group at the beginning of the period.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.4: Labor demand adjustments: new hires

	All (1)	Young (2)	Middle-aged (3)	Old (4)	Permanent (5)	Temporary (6)
$T_i \cdot$ Post	-0.074**	-0.042**	-0.032**	-0.000	-0.002	-0.072**
	(0.034)	(0.019)	(0.015)	(0.004)	(0.014)	(0.028)
$R_i \cdot$ Post	-0.348**	-0.197**	-0.149**	-0.002	-0.011	-0.337**
	(0.159)	(0.090)	(0.072)	(0.018)	(0.067)	(0.131)
Mean outcome pre-2012	4.79	2.38	2.06	0.35	1.46	3.32
Observations	430,038	430,038	430,038	430,038	430,038	430,038
Mean T_i	1.37	1.37	1.37	1.37	1.37	1.37
Std. Dev. T_i	1.33	1.33	1.33	1.33	1.33	1.33
Mean R_i	0.83	0.83	0.83	0.83	0.83	0.83
Std. Dev. R_i	0.99	0.99	0.99	0.99	0.99	0.99
1st stage coeff.	0.21***	0.21***	0.21***	0.21***	0.21***	0.21***
1st stage F-stat.	8,288.71	8,288.71	8,288.71	8,288.71	8,288.71	8,288.71

Notes: The table reports estimates from specification (1.4), which includes firm and year fixed effects. Standard errors in parentheses are clustered at the firm level. Column 1 shows the effect on new hires of all workers, columns 2 to 4 on young ($\leq$ 35 years old), middle-aged (aged 35-55) and old (above 55 years old) workers, respectively. Columns 5 and 6 show the effect on workers hired under permanent and temporary contracts, respectively. In the first row the treatment is T_i, the average change in the full retirement date of *potential retirees* employed at the firm when the pension reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. In the second row the treatment is R_i, the number of *potential retirees* who retire at least one year after the predicted pre-reform full retirement date, and T_i is used as an instrument. The first stage F statistics in the IV regression is the Kleibergen-Paap rk Wald F statistics. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.5: Incumbents' labour earnings

	Co-workers					*Potential Retirees*
	All	All (w/ non-work subsidies)	Permanent employees Young	Middle-aged	Old	All
	(1)	(2)	(3)	(4)	(5)	(6)
$T_i \cdot$ Post	-6,650.7***	-4,902.0***	-1,220***	-4,567.3***	-1,520.9***	5,107***
	(1,787.5)	(1,569.9)	(514.6)	(1,201.5)	(316.3)	(251.5)
Observations	423,346	423,346	419,363	419,363	419,363	425,971
Mean outcome pre-2012	647,005.96	648,244.1	138,497.38	440,187.65	42,319.77	49,951.32
Mean T_i	1.36	1.36	1.36	1.36	1.36	1.36
Std. Dev. T_i	1.32	1.32	1.31	1.31	1.31	1.33

Notes: The table shows the response of labour earnings of *potential retirees* and of their full-time co-workers to a 1-year increase in T_i, the shift in the full retirement date of *potential retirees* employed at the firm when the reform is passed. Standard errors in parentheses are clustered at the firm level. Worker-level observations are aggregated at the level of the firm where workers were incumbents in the last quarter of 2011. Column 1 and 3-6 do not include non-work subsidies, while column 2 does. Earnings and non-work subsidies are winsorized at the 99th and 99.99th percentiles, respectively. Columns 3 to 5 display the effect on permanent young ($\leq$ 35), middle-aged (36-55), and older (over 55) co-workers, respectively. The regressions are based on specification (1.4) and are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. Observations are weighted according to firm size at the beginning of the period.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.6: Fiscal Externality

	Potential retirees (1)	All $\tau = 30$ (2)	All $\tau = 35$ (3)	All $\tau = 40$ (4)	W/ Loss on Non-Hired (5)
$P^e = 0.7 \cdot P^{med}$	0.334	-0.567	-0.592	-0.621	-0.656
	(0.020)	(0.011)	(0.124)	(0.143)	(0.124)
$P^e = 0.7 \cdot P^{mean}$	0.327	-0.597	-0.624	-0.655	-0.689
	(0.020)	(0.109)	(0.126)	(0.146)	(0.126)
$P^e = 0.9 \cdot P^{med}$	0.291	-0.611	-0.637	-0.667	-0.70
	(0.020)	(0.105)	(0.122)	(0.141)	(0.122)

Notes: The table reports estimates of the fiscal externality based on formula (1.7). A negative externality between -1 and 0 implies that behavioral responses generate costs, but they are smaller than savings on pension outlays. A positive fiscal externality implies that behavioral responses generate additional resources for the government on top of mechanical savings on pension spending. The first row calibrates P^e by using the median pension (P^{med},13,127 euros); the second uses the mean pension (P^{mean}, 16,279 euros); the third uses the median pension and calibrates $P^e = 0.9 \cdot P^{med}$, while $P^e = 0.7 \cdot P^{med}$ and $P^e = 0.7 \cdot P^{mean}$ in the first and the second row, respectively. Column 1 reports estimates that ignore the spillover on co-workers and sets $\tau = 35$ percent. Columns 2 to 4 show calibrations where all spillovers are included, across alternative calibrations of τ. Column 5 reports estimates from a model that augments formula (1.7) assuming that every marginally non-hired worker earns zero labor earnings for as long as the median duration of job search for workers who find a job over the 2012-2015 period (3 months). We calibrate the foregone earnings by using an estimation of the median value of labor income earned in the first 3 months after starting a job and we employ estimates on the effect of the treatment of new hires to calibrate the number of marginally non-hired workers. In this column, the tax rate is calibrated as $\tau = 35$ percent. More details on calibrations and estimation are reported in Appendix Section 1.G.

Appendices

1.A The Italian labour market

Firm size: Italy is the European country that features the highest number of enterprises, totalling around 3.8 millions in 2012.[56] 95 per cent of Italian firms are considered micro-enterprises and have less than 9 employees. The share of employees in firms with less than 250 employees was around 78.7 per cent in 2015, compared to 62.9 per cent in Germany and 61.4 per cent in France.[57] As we conduct our analysis on firms which have between 3 and 200 employees in the first quarter of the sample period, we are considering a sample that is highly representative of the Italian productive landscape.

Workforce demography: The age structure of the Italian workforce underwent profound changes during the last decade. The share of workers aged between 55 and 64 increased from 10 per cent in 2005 to 16.8 per cent in 2015.[58] France and Germany experienced similar trends, with a 4.9 percentage point (from 10.4 per cent) and a 6.4 percentage point (from 12.4) increase, respectively. Understanding the consequences of retaining older workers at firms is therefore of great relevance.

Employment protection legislation: Italy is one of the countries with the highest degree of employment protection in Europe, together with Germany and France.[59] Two labour reforms - the 2012 *Fornero* reform (Law 92/2012) and the 2015 *Jobs Act* (Decree Law 23/2015) - have however progressively reduced the costs that firms bear in case of unfair dismissals, while increasing their predictability. Before 2012 a permanent worker who was found to be unfairly dismissed from a firm with more than 15 employees had the right to be reinstated in the firm. The labour reforms have substantially narrowed the circumstances that lead to the reinstatement of a worker; they have also limited judges' discretion in setting the amount of monetary compensation. For workers hired with an open-ended contract after March 7, 2015 severance payments in case of unfair dismissal were legislated to be a

[56] Data from Eurostat, Annual Enterprise Statistics. Financial and insurance sectors are excluded.

[57] Figures are the result of authors' computations based on Eurostat Annual Enterprise Statistics and National Accounts Statistics.

[58] Source: Eurostat, Employment Statistics.

[59] See OECD (2015) data on employment protection legislation.

deterministic and increasing function of tenure: 2 months of the last salary per year of service, with a minimum of 4 months and a maximum of 24 months. The 2018 labour reform (Decree Law 87/2018) has raised the minimum and the maximum to 6 and 36 months, respectively. Furthermore, a sentence of the Constitutional Court has deemed unconstitutional that the indemnity in case of unfair dismissals is only a function of tenure.

1.B Additional details about the *Fornero* reform

Grandfathering clauses: The new rules introduced by the *Fornero* reform do not apply to individuals who satisfied the requirements for claiming either an old-age or a seniority pension by the end of 2011. It was furthermore legislated that some specific categories of employees were grandfathered as well. These are mainly workers who, during the passage of the reform, were collocated on redundancy schemes or on short-time work programs. According to the law, the categories of private-sector employees who could still retire under pre-reform rules are therefore the following:

i) Workers who satisfy the requirements for claiming either an old-age or a seniority pension by 12/31/2011;

ii) Workers *collocati in mobilitá* according to law 223/91 and based on collective agreements signed before 12/4/2011. Workers *collocati in mobilitá* were laid-off workers who received specific monetary support and were engaged in redeployment programs;

iii) Workers who, as of 12/4/2011, were beneficiaries of *prestazioni straordinarie a carico dei fondi di solidarietá di settore.* These are workers in short-time work who received monetary support from *ad-hoc* sectoral solidarity funds;

iv) Workers who, as of 12/4/2011, had ceased to work but had been authorized to continue paying contributions.

v) Workers who, as of 12/4/2011, had been granted the *istituto di esonero dal servizio.*

In the following years, further specific categories of workers where grandfathered (the so-called *salvaguardati*). In our analysis, when predicting full retirement dates under pre- and post-reform rules, we only apply grandfathering clauses to workers listed in (i), because we do not have the necessary information to identify employees belonging to categories (ii) to (v). Furthermore, as far as *salvaguardati* are concerned, we do not take into account changes to the law that occurred in later periods, because we aim to build a measure of the shock to the full retirement date at the time when the reform came into force.

Other provisions: Beyond the changes shown in Table 1.1 and Appendix Table 1.H.1, the *Fornero* pension reform also introduced the following provisions, which we take into account when computing the predicted full retirement dates as described in Subsection 1.5.1.1:

i) Women who were at least 60 years old and had at least 20 years of contributions by 2012 can exceptionally retire upon turning 64 years old in 2012, 64 and 3 months old in 2013-2015 and 64 years and 7 months old from 2016 onward. This possibility is granted to all workers who would have reached "quota 96" in 2012.

ii) Before 2012, individuals who started working in 1996 or later could also claim an old-age pension upon satisfying the same age requirement as other individuals, but it was conditional on a lower contribution requirement: 5 years of *effective* contributions, rather than 20. The *Fornero* reform maintains the same milder contribution requirement, but substantially raises the age requirement: 70 years old in 2012, 70 years and 3 months old in 2013-2015 and 70 years and 7 months old from 2016 onward.

iii) Workers who started working after 1995 can also claim a seniority pension upon turning 63 years old in 2012, 63 years and 3 months old in 2013-2015 and 63 years and 7 months old from 2016 onward, which is conditional on having at least 20 years of contributions.

1.C Procedure to clean matched employer-employee data

Firm covariates and outcomes come from matched employer-employee data over the period 2009-2015. The unit of observation is the worker-firm relationship in a given month. More than one relationship between a worker and firm in a given month may exist. This is because firms are required to compile two *Uniemens* modules for a given employee if a characteristics of his/her contract changes during the month. In such a case, we isolate and retain only the prevailing relationship, according to the following multi-step procedure:

i) We drop records where the wage is set equal to 0. If all records have this feature, we keep one randomly.

ii) If there are records that feature the same contract characteristics (occupation, duration, full-time or part-time status, type of collective contract) and the same wage, we drop all but one randomly.

iii) We drop records that feature lower numbers of paid days.

iv) When multiple records arise in a given month, but not in the months immediately before or after that, we look at the characteristics of the worker-firm relationship in the preceding and in the following month. We then keep the single record that satisfies the following (ranked) criteria: a) modal occupation; b) wage closest to the average one in the neighbouring months; c) highest number of paid days; d) highest wage.[60] If more than one record survives criteria (a) to (d), we drop all but one randomly.

v) When multiple records arise in each of a set of consecutive months, within each month we keep the single record that satisfies the following (ranked) criteria: a) highest number of paid days; b) highest wages.[61] If more than one record survives criteria (a) to (b), we drop all but one randomly.

[60] If more than one record satisfies criterion (a), we then use criterion (b), and so on up to criterion (d).

[61] If more than one record satisfies criterion (a), we then use criterion (b).

1.D Computation of years of qualifying contributions

Contributions are of two types: *effective* contributions, which arise as a result of periods of paid work, and *figurative* contributions, which arise as a result of events that include sick leave, maternity leave, short-time work, unemployment and disability. *Figurative* contributions are not paid out by the workers, but they nevertheless accrue on their accounts. Depending on the type of pension, *figurative* contributions may not count toward the accrual of the right to retire (while still counting toward the determination of the amount of the pension benefit).

Workers' contribution histories record the event giving raise to each contribution spell, allowing us to distinguish *effective* contributions from *figurative* ones. For every type of pension, we therefore only take into account relevant contributions, improving the accuracy of predicted retirement dates. We first sum up contribution spells (expressed in weeks) in any given year, capping them at 52 weeks, which is the maximum number of weeks of contributions acquirable every year. Following rules for totalling contributions used at INPS, in case of (partially or totally) overlapping spells we avoid double counting. We then sum up contributions across years, up to December 2011. The underlying assumption is that, in case of workers who accrue contributions across different funds, they choose to (onerously) exercise the so-called *ricongiunzione* option, which allows them to bring all contributions together into a unique fund, so that they can be summed up toward the accrual of pension rights.

1.E Conceptual framework

To guide our empirical analysis of firms' responses to pension reforms, we outline a labour demand model that features a shock to the retention rate of older workers. We focus on firm-driven changes in the employment of their co-workers. We then investigate how this response relates to the degree of substitutability between *potential retirees* and their co-workers. We start by analyzing a standard model where we remain agnostic about the wage formation process. We then study the behaviour of labour demand in different wage bargaining settings. First, we analyze the standard Nash-bargaining model. Second, we introduce Nash-bargaining over profits to capture the profit-sharing behaviour that has been documented by Card et al. (2013) in the Italian context. Third, we study a monopsonistic labour market with constant labour supply elasticity. Consistently across settings, the change in labour demand is inversely proportional to the degree of substitutability between *potential retirees* and co-workers.

1.E.1 Labour demand model

Consider a two-period model where the firm chooses the optimal employment in period 1 given the employment in period 0. We assume that there are two types of workers: *potential retirees* (p) and co-workers (c). In our empirical setting *potential retirees* are older workers close to retirement, and co-workers are other older workers further away from retirement or younger employees in the same firm. Denote with n_0^c and n_1^c the number of co-workers employed in period 0 and 1, respectively. Adjustments in the demand for co-workers are referred to as x^c, so that $n_1^c = n_0^c + x^c$. A cost function $c\left(x^c\right)$ accounts for the cost - paid in period 0 - of adjusting the co-workers workforce. We require $c\left(\cdot\right)$ to be twice continuously differentiable and we assume that $c'\left(x^c\right) > 0$ for $x^c > 0$, $c'\left(x^c\right) \leq 0$ for $x^c < 0$, $c'\left(0\right) = 0$ and $c''\left(\cdot\right) \geq 0$. This cost function is flexible enough to incorporate any asymmetry in adjusting downwards or upwards the co-workers' labour demand. For the sake of simplicity, we assume that no *potential retiree* can be either hired or fired. We denote with n_0^p and $n_1^p = sn_0^p$ the number of *potential retirees* in period 0 and 1, respectively. $s \leq 1$ captures the exogenous share of *potential retirees* who are left in period 1. We interpret s as a variable incorporating the exogenous separation rate of *potential retirees* as well as retirement rules. Output is produced according to technology $F\left(n_t^p, n_t^c\right)$ in every period $t = 0, 1$, with $F_p, F_c \geq 0$, $F_{pp}, F_{cc} \leq 0$, and we impose no restriction on cross derivatives. The firm is wage and price taker, and the price of output is normalized to 1. The demand for co-workers in period 1 is chosen so as to maximize profits, which are given by

$$\pi = \pi_0 + \beta\left(F\left(sn_0^p, n_0^c + x^c\right) - w^p sn_0^p - w^c\left(n_0^c + x^c\right)\right) - c\left(x^c\right) \tag{1.8}$$

where π_0 are profits in period 0, β is a discount factor, and w^p and w^c are the wages in period 1 of *potential retirees* and co-workers respectively. Optimality conditions require the following

$$\beta \left(F_c \left(s n_0^p, n_1^c \right) - w^c \right) = c' \left(x^c \right) \tag{1.9}$$

The firm equates the marginal increase in revenues net of wage expenditures to the marginal cost of adjusting co-workers' labour demand. A change in retirement rules that increases the retirement age can be approximated by a smaller than expected drop in the number of *potential retirees* in period 1, *i.e.* an increase in s. The comparative statics for a change in s is

$$\frac{\partial x^c}{\partial s} \propto \beta \left(F_{pc} n_0^p - \frac{\partial w^c}{\partial s} \right) \tag{1.10}$$

The sign of the comparative statics depends on two terms. First, the degree of substitutability between the two types of labour captured by F_{pc}. Second, the extent to which wages adjust after the policy shock. If the two types of workers are substitutes - that is if $F_{pc} < 0$ - only a strong decrease in w^c can lead to an increase in the demand for co-workers. Indeed, in order to hire co-workers, the firm must significantly cut the payroll to compensate the loss in marginal productivity of co-workers that follows an exogenous increase in the number of *potential retirees*. However, wages are usually expected to be sticky, with the implication that when the two types of workers are substitutes we likely observe a drop in the demand of co-workers. We present here a few interesting cases. First, if wages are sticky (*i.e.* $\partial w^c / \partial s = 0$), the response of co-workers' labour demand depends on the substitutability between co-workers and *potential retirees*; if the two types are substitutes the firm decreases demand for co-workers. Second, if wages are flexible and partially follow the change in the marginal productivity of co-workers (i.e. $\partial w^c / \partial s = \alpha n_0^p F_{pc}$ with $\alpha < 1$), labour demand decreases as long as the two types of work are substitutes.[62] Finally, in a competitive labour market where wages reflect the marginal productivity of co-workers we would have no change in labour demand since prices fully adjust to absorb the shock.

Result 1: *Evidence of a drop in labour demand for co-workers can be reconciled with complementarity between co-workers and potential retirees only in case of a large increase in the wages of co-workers.*

We document in Section 1.7 that the pension reform causes a drop in labour demand for co-workers. Moreover, a large increase in co-workers' wages is inconsistent with the evidence we provide in Section 1.9, which shows a drop in earnings for

[62] There are different explanations for having co-workers' wages non perfectly reflecting their marginal productivity. Lazear (1979) shows in a dynamic model that an increasing wage path where older workers are overpaid can be used to provide incentives to young workers.

co-workers. We conclude that co-workers are substitutes for *potential retirees*.

1.E.2 Alternative wage formation models

1.E.2.1 Intrafirm bargaining over total surplus

So far we have been agnostic about the wage formation process. We now consider the case where wages are set according to Nash bargaining between the firm and individual workers over the total surplus of the match, as is standard in the labour search literature. Suppose that co-workers have bargaining power ϕ and outside option $\underline{w}^c$. The surplus generated by a match is a function of the marginal profit generated by the worker and of the wedge between the wage and the worker's outside option. We allow all wages to be re-negotiated in period 1. In equilibrium:

$$\phi \frac{\partial \pi(sn_0^p, n_1^c)}{\partial n_1^c} = (1-\phi)(w^c - \underline{w}^c) \tag{1.11}$$

which implies the following expression for the equilibrium wage:

$$w^c = \underbrace{\eta F_c - \frac{\eta}{\beta} c'(x^c)}_{\text{Marginal output net of adjustment costs}} + \underbrace{\frac{(1-\phi)}{\phi\beta} \eta \, \underline{w}^c}_{\text{Reservation wage}} \tag{1.12}$$

where $\eta = \phi\beta/(\phi\beta + 1 - \phi)$. When co-workers have no power in the bargaining the wage is set exactly equal to the outside option. The expression is analogous to the one derived by Cahuc et al. (2008). Co-workers' wages in equilibrium are a function of their marginal output net of marginal cost and of their reservation wage. We are interested in the effect of a change in s on wages that reads:

$$\frac{\partial w^c}{\partial s} = \eta[F_{pc} n_0^p + F_{cc} \frac{\partial x^c}{\partial s}] - \frac{\eta}{\beta} c''(x^c) \frac{\partial x_1^c}{\partial s} \tag{1.13}$$

The wage change in response to a shock to the retention rate depends on the cross-marginal product between co-workers and *potential retirees* (F_{pc}), as well as on the slope of co-workers' marginal product (F_{cc}). The last term in (1.13) arises since we do not assume linear adjustment costs. Notice that we implicitly relied on the assumption that the worker's outside option does not change per effect of the reform. This is because we consider - as in our empirical implementation - a firm-specific shock to the retirement age. The assumption would be violated if the general equilibrium effects of the reform were large.

By using (1.13) in (1.10) we get the following expression for the adjustment in labour demand of co-workers in period 1:

$$\frac{\partial x^c}{\partial s} = -\frac{\beta F_{pc} n_0^p}{\beta F_{pp} - c''(x^c)} \tag{1.14}$$

When the reservation wage does not change, the shift in s does not have any first order effect on the wage. Hence, there is a one to one mapping between workers' complementarity and the change in labour demand.

Result 2: *In a model of intra-firm bargaining where workers and firms bargain over marginal profits and worker surplus, there is a one-to-one relationship between changes in the labour demand of co-workers and the complementarity between the two types of labour. It follows that a drop in co-workers' labour demand caused by a change in s is only consistent with substitutability between potential retirees and co-workers.*

1.E.2.2 Intrafirm bargaining over profits

Card et al. (2013) present evidence of substantial profit sharing in Italian firms. We extend our model to account for profit sharing by allowing firms and workers to bargain over total profits and co-worker's surplus. In equilibrium:

$$\phi\pi = (1-\phi)\left(w^c - \underline{w}^c\right) \tag{1.15}$$

This implies the following:

$$\left(1-\phi+\beta\phi n_1^c\right) w^c = \beta\phi\left(\frac{\pi_0}{\beta} + \left(F - w^p s n_0^p\right) - \frac{1}{\beta} c\left(x^c\right)\right) + (1-\phi)\,\underline{w}^c \tag{1.16}$$

Wages are determined by profits net of co-workers' costs and by workers' outside option. We totally differentiate equation (1.16) assuming that the wage of *potential retirees* is sticky to find an expression for the wage response to a change in s:

$$\frac{\partial w^c}{\partial s} = \tilde{\eta}\left(F_p - w^p\right) \tag{1.17}$$

where $\tilde{\eta} = (\beta\phi n_0^p)/(1-\phi+\beta\phi n_1^c)$. Because of an envelope argument, the effect of the reform on co-workers' wages is proportional to the wedge between *potential retirees*' productivity and wage. Intuitively, a larger gap increases the passthrough of the reform to co-workers' wages. In response to the change in profits caused by the reform, firms decrease the salary of co-workers to preserve the wedge for *potential retirees*.[63] After replacing (1.17) in (1.10) it follows that if wages for co-workers decline, labour demand can drop only in the case $F_{pc} < 0$.

[63]If firms where able to adjust *potential retirees*' wages the total passthrough on co-workers would be smaller.

Result 3: *In the case where potential retirees get paid more than their productivity, the reform causes a drop in co-workers' salaries. Therefore, evidence of a fall in co-workers' labour demand can only be reconciled with substitutability between co-workers and potential retirees.*

1.E.2.3 Monopsonistic labour market

We consider the broadly used model of monopsonistic labour demand and we solve a simple version with constant elasticity of labour supply. Suppose the firm was not a wage taker and chose employment knowing the labour supply elasticity and anticipating the consequences of labour demand on the wage. We further assume that the labour supply of co-workers is such that $n_1^c = w^e$, where e is the elasticity of labour supply to the wage and $e > 0$. The firm's problem would become:

$$\pi = \pi_0 + \beta \left(F\left(sn_0^p, n_1^c\right) - w^p s n_0^p - n_1^{c\frac{1+e}{e}} \right) - c\left(x^c\right) \tag{1.18}$$

The firm's optimality condition in this model is:

$$\beta \left(F_c - \frac{1+e}{e} n_1^{c\frac{1}{e}} \right) = c'\left(x^c\right) \tag{1.19}$$

From 1.19 we derive a new comparative statics:

$$\frac{\partial x^c}{\partial s} = - \frac{F_{pc} n_0^p}{F_{cc} + \frac{1+e}{e^2} n_1^{c\frac{1-e}{e}} - \frac{1}{\beta} c''(x^c)} \tag{1.20}$$

The expression above shows a one-to-one mapping between labour demand changes and the substitutability between *potential retirees* and co-workers. The extent to which labour demand drops if $F_{pc} < 0$ decreases with the elasticity of labour supply. When labour supply is more elastic, the firm has limited room to adjust co-workers' labour demand in response to the reform.

Result 4: *A monopsonistic labour market delivers a one-to-one relationship between co-workers' labour demand responses and the substitutability between co-workers and potential retirees. If the two types of work are substitutable, co-workers' labour demand falls in response to a positive shock to the retention rate of potential retirees.*

1.F Matching procedure and the cost of layoffs

Matching procedure and estimation: Matching covariates measured in 2009 are: age, sex, wage, occupation, a dummy for permanent contracts, a dummy for full-time contracts, experience, sector, province and firm size. We partition each continuous variable in several bins and match only control workers who fall in the same combination of bins as at least one laid-off worker. We call this combination a strata. After we match fired workers to workers who do not separate from the firm where they were incumbent in the last quarter of 2011, we estimate the following specification:

$$Y_{it} = \lambda_i + \sum_{k=-2}^{3} \beta_k \gamma_k + \sum_{k=-2}^{3} \beta_k^l \gamma_k \cdot \text{Layoff}_i + \sum_{k=-2}^{3} \beta_k^x \gamma_k \cdot X_i + \varepsilon_{it} \tag{1.21}$$

X_i is the vector of matching covariates and $k = 0$ is the event year in which the layoff takes place. We focus on layoffs occurring in years 2012 and 2013. Since our sample ends in 2015 we estimate a model with only 3 periods after the layoff to make sure that all coefficients are identified by the same number of observations. We then impute the estimate of β_3^l in (1.21) as the cost of job losses four years after the layoff event. The regression is weighted based on the weights defined below.

Coarsened Exact Matching (CEM) weights: Let N_C and N_T be the number of control and treatment units in the matched sample. Suppose we have S strata where $s = 1, ..., S$ and each of them contains N_{Ts} treated unit and N_{Cs} control units. The CEM weight for a control unit is the following:

$$w_i = \frac{N_C}{N_T} \cdot \frac{N_{Ts}}{N_{Cs}}$$

while each treated unit receives weight equal to 1 (see Iacus et al., 2012). This guarantees that weights sum up to total matched observations:

$$\begin{aligned}
\sum_i w_i = \sum_{i \in C} w_i + \sum_{i \in T} w_i &= \sum_{i \in C} w_i + N_T \\
&= \frac{N_C}{N_T} \sum_s \sum_{i \in s} \frac{N_{Ts}}{N_{Cs}} + N_T \\
&= N_C + N_T
\end{aligned}$$

1.G The Fiscal Externality

1.G.1 Derivation of The Fiscal Externality

We consider two types of agents and we define them as in our empirical analysis with the labels of *potential retirees* (p) and co-workers (c). Agents perform different labour-related activities. We call l_i^j the share of individuals of type i performing activity j. Each agent faces the following budget constraint:

$$\begin{aligned} x_i \leq & \left(1-\tau_i\right) w_i l_i^w + \left(1-\tau_i\right)\left(NW_i l_i^{NW} + D_i l_i^D\right) \\ & + \left(1-\tau_i^P\right)\left(P_i\left(\bar{T}-T_i^P\right)\cdot \mathbb{1}\left(\bar{T}>T_i^P\right)\cdot \mathbb{1}\left(l_i^E=0\right) + P_i^E l_i^E\right) + y_i \end{aligned} \tag{1.22}$$

where $\left\{\tau_i, \tau_i^P, NW_i, D_i, P_i, T_i^P, P_i^E\right\}$ is a vector of policies targeting agent i. τ_i is an average labour earnings tax rate, τ_i^P is an average tax rate on pension payments, NW_i are non-work subsidies, D_i are disability benefits, P_i are regular pension entitlements, T_i^P is the full retirement date, and P_i^E are pension benefits for workers who retire early.[64] $\bar{T}$ is our evaluation horizon. w_i denotes the wage, $z_i = w_i l_i$ is total labour earnings, and y_i is income not related to working activities. We model the reform as a change in the full retirement date T_i^P. If after an increase in T_i^P a worker retires at the previously expected date, she will receive a lower pension payment because $P_i^E < P_i$.

The fiscal externality of the policy is the share of mechanical revenues that is lost because of the behavioural responses:

$$FE = -\frac{\sum_{i=p,c} n_i \left[\left(1-\tau_i\right)\left(NW_i \frac{dl_i^{NW}}{dT_i^P} + D_i \frac{dl_i^D}{dT_i^P}\right) + \left(1-\tau_i^P\right) P_i^E \frac{dl_i^E}{dT_i^P} - \tau_i \frac{dz_i}{dT_i^P}\right]}{\sum_{i=p,c} n_i dT_i^P \left(1-\tau_i^p\right) P_i \cdot \mathbb{1}\left(\bar{T}>T_i^P\right)\cdot \mathbb{1}\left(l_i^E=0\right)} \tag{1.23}$$

where n_p and n_c denote the number of *potential retirees* and co-workers, respectively. The numerator represents the costs incurred by the government because of behavioural responses. Mechanical revenues in the denominator instead measure the resources that the government would save through the policy from incumbent employees, absent any change in the behaviour of workers and firms.

1.G.2 Empirical implementation and results

The fiscal externality is a function of the estimates in Sections 1.7 and 1.9.[65] The terms referring to NW and D in the numerator of equation (1.23) measure the budget consequences of the reform on other policy instruments. We quantify them using causal estimates of the effect of the reform on these outcomes. The last term in the numerator of (1.23) is the effect on labour income tax revenues. It is a function of

[64] Notice that when workers retire early they do not receive the full pension payment P_i.
[65] Coefficients are reported in Appendix Table 1.H.15.

the causal effect of the reform on *potential retirees* and co-workers' earnings. Finally, the term $P_i^E \frac{dl_i^E}{dT_i^P}$ measures the impact of changing the full retirement age on early retirement. To quantify it we need $\frac{dl_i^E}{dT_i^P}$, which we get by estimating the effect of the reform on months spent in retirement before the statutory retirement date. We calibrate P^E as a conservative 70 per cent of the average or median value of yearly pension payments in the data (roughly 13,100 and 16,300 euros, respectively).[66] We check alternative parametrizations of τ ranging from 30 per cent to 40 per cent for robustness. Notice that the average personal income tax rate for the median income (roughly 22,000 euros) is 24 per cent without considering tax credits and social security contributions. We calibrate τ_i^P starting from τ_i and including the tax credit available for the median or average value of the pension payment, depending on the one we use in the calibration.[67] Finally, we obtain the mechanical effects in the denominator of (1.23) by subtracting the behavioural response $P_i^E \frac{dl_i^E}{dT_i^P}$ to causal estimates of the effect of the policy on pension outlays.

To calibrate the foregone earnings of individuals who are not hired in column 5 of Table 1.6 we employ publicly available Labor Force Survey data provided by the Italian National Institute of Statistics (ISTAT). We calibrate the median labour earnings losses using the median labour earnings of new hires in the 3-month period following the hiring event.[68] First, we compute the share of new hired workers for any combination of full-time/part-time and temporary/permanent contracts. Second, we compute the median number of hours worked for each of these groups. Third, we compute the median contract duration for temporary contracts to assess for which portion of the 3-month period the workers would be employed in the counterfactual scenario where they would find a temporary contract job. Since the median duration of temporary contracts is above 3 months, we compute the losses over the entire 3-month period for both permanent and temporary contracts. We then combine all these calibrated values with a hourly wage of 9.89 euros (ISTAT).[69] Our estimate for the earnings loss is 2996.18 euros for every worker who is not marginally hired.

---

[66] Workers claiming *opzione donna* (the main early retirement option) get on average roughly 65 per cent of full pension benefits that an individual would be entitled to if choosing the seniority retirement option (INPS, 2016). Also, a small number of workers can retire before the statutory date obtaining full pension entitlements thanks to some grandfathering clauses (see Appendix Section 1.B). Hence, in the latter case our calibration understates the benefit received when employees retire before the post-reform statutory date. Therefore, we also show a calibration whereby $P^e = 0.9 \cdot P$.

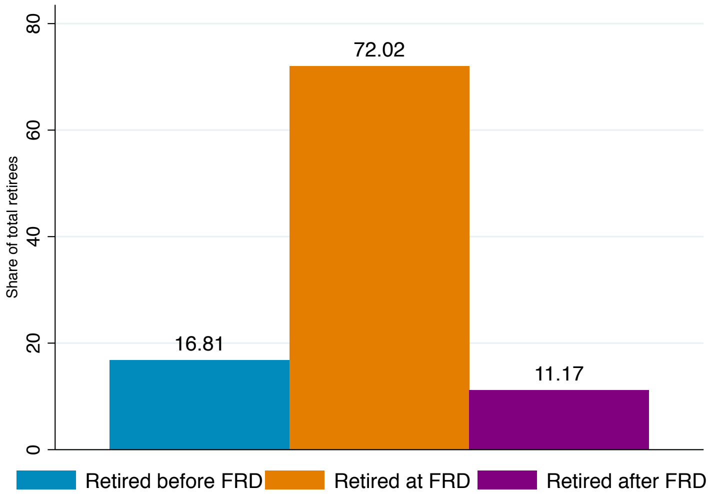
[67] For pensions below 15,000 euros the tax credit is equal to $1297 + (583 \times (15000 - \text{Pension})/7000)$. For pensions between 15,000 and 55,000 euros the formula is $1297 \times (55000 - \text{Pension})/40000$.

[68] 3 months is the average job search period for workers who find a job in 2012-2015.

[69] The number is the average of the average hourly wage for new hires in the period 2014-2016. It is the result of (9.74+9.95+9.99)/3 (ISTAT, *I Differenziali Retributivi nel Settore Privato*).

1.H Additional figures and tables

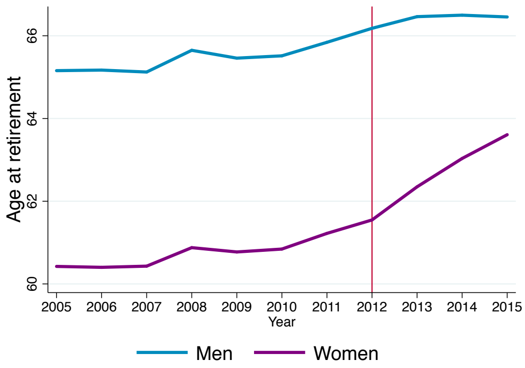
Figure 1.H.1: Share of workers retiring at the full retirement date



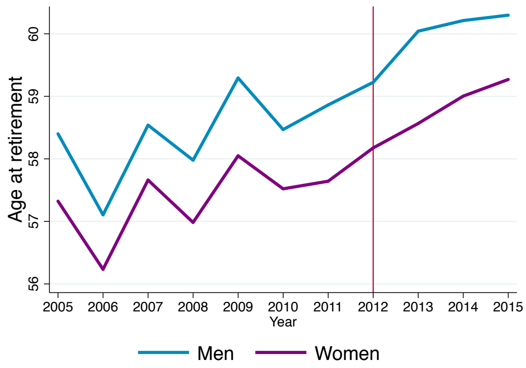
Notes: The figure shows the share of individuals who: (i) retire in a 1-year window around the earliest among their full retirement dates ("Retired at FRD"); (ii) retire more than 1 year before reaching the earliest among their full retirement dates ("Retired before FRD"); (iii) retire more than 1 year after reaching the earliest among their full retirement dates ("Retired after FRD").The full retirement dates are the dates when an individual becomes eligible to receive benefits under: (i) the old-age pension scheme; (ii) the seniority pension scheme. We define an individual as retired when he/she starts collecting Social Security benefits. Shares are derived from authors calculations on the INPS register of retirees.

Figure 1.H.2: Age at retirement and number of new retirees by gender and type of pension

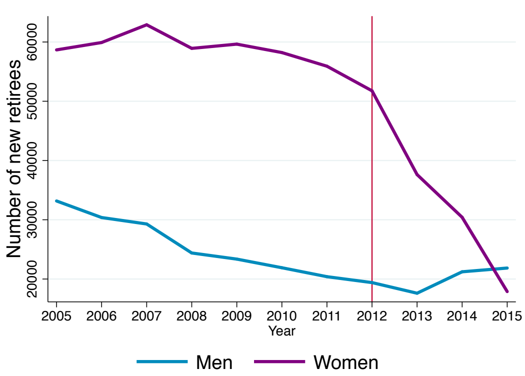
(a) Old-age pensions



(b) Seniority pensions



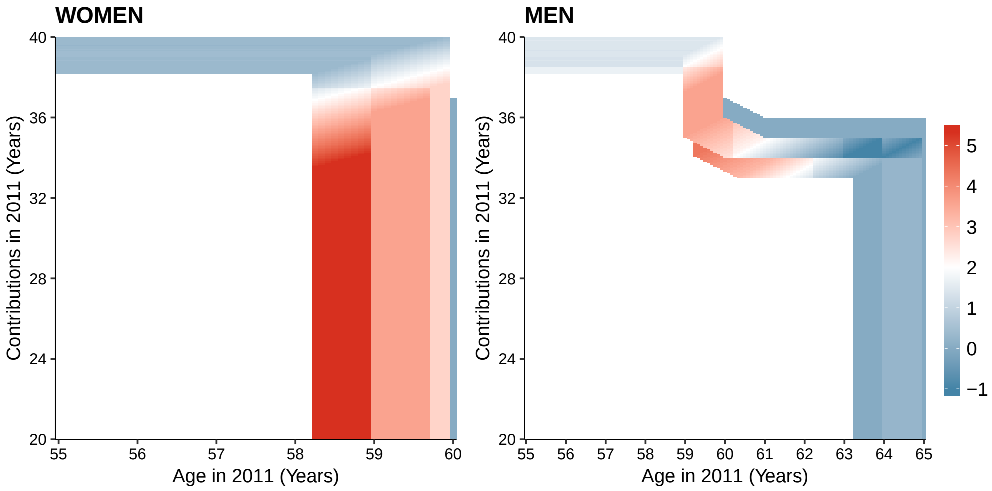
(c) Old-age pensions



(d) Seniority pensions

Notes: Panels (a) and (b) show the evolution of age at retirement of new retirees who worked as private-sector employees, split by gender and type of pension. Panels (c) and (d) plot the evolution of the two-year moving average (using lags only) of the number of new retirees who worked as private-sector employees, split by gender and type of pension. We define an individual as retired when he/she starts collecting retirement benefits. The vertical line represents the year when the *Fornero* pension reform becomes effective (2012).

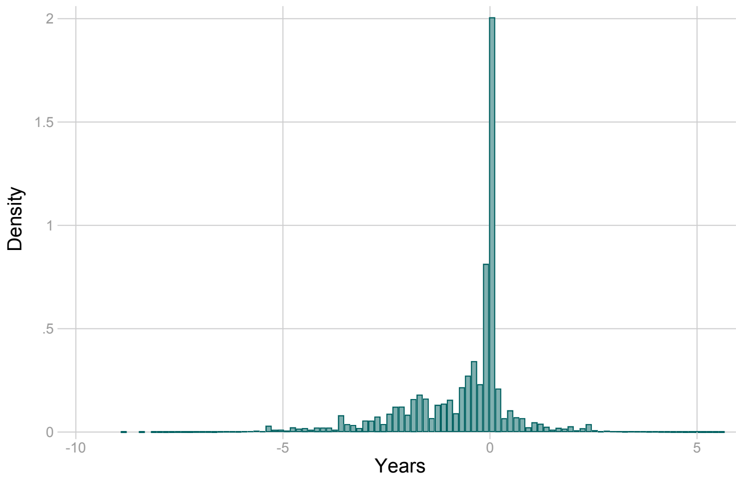
Figure 1.H.3: Reform-induced changes in the full retirement date



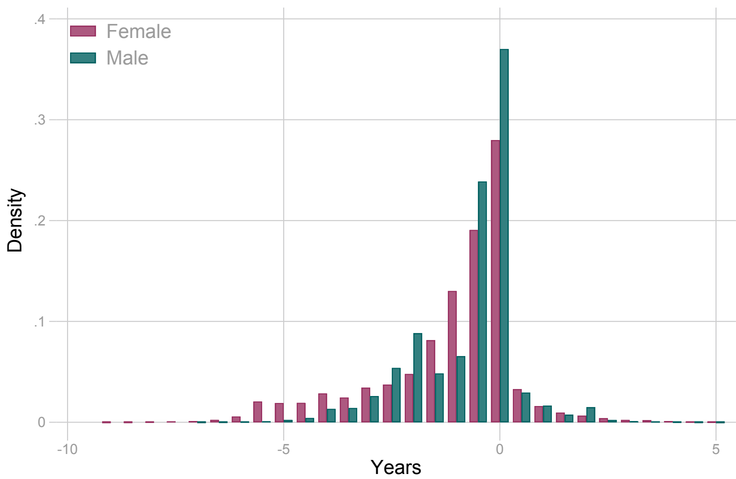
Notes: The figure plots heat-maps showing the relationship between the reform-induced change in the retirement date and the characteristics of the worker in 2011. The characteristics are workers' gender, age, and years of contributions in December 2011. The changes are constructed using the rules detailed in Table 1.H.1 under the assumptions listed in Section 1.5.1.1. The figure only reports changes for workers expected to start collecting pension benefits by 2014 at the reform date. The blank areas in the bottom left of the two plots are combinations of age and years of contributions that do not ensure eligibility to collect retirement benefits by 2014. The blank areas in the top right of the two plots are combinations of age and years of contributions that ensure eligibility to collect retirement benefits in 2011 or before.

Figure 1.H.4: Post-reform retirement date - Forecast quality assessment

(a) All workers



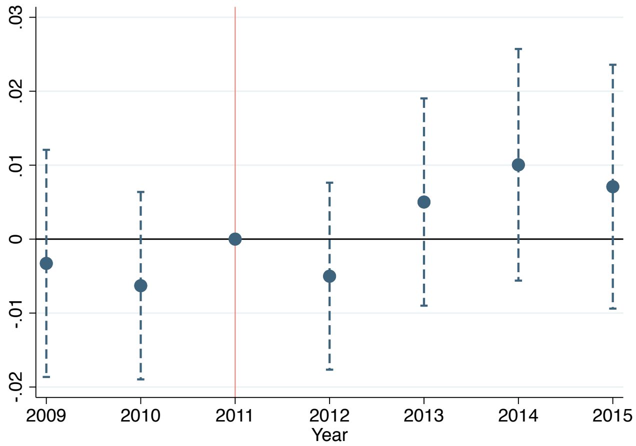
(b) By gender



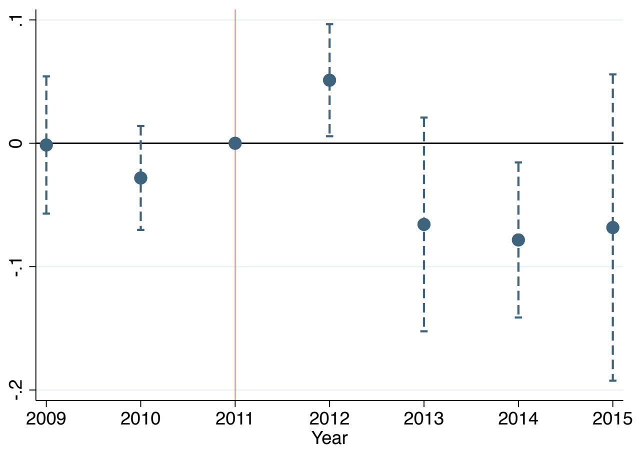
Notes: The figure shows a forecast quality assessment of the prediction of the post-reform retirement date. The horizontal axis measures the difference (in years) between the observed retirement date and the predicted full retirement date under post-reform rules. The sample includes all workers who were expected to retire by 2014 under pre-reform rules and retired in the period 2012-2017. A positive difference implies that a worker retires after his/her predicted full retirement date, a negative difference means that the individual retires earlier than predicted. We define an individual as retired when he/she starts collecting Social Security benefits. Panel (a) plots the distribution for the entire sample of workers, Panel (b) shows the breakdown by gender. Number of workers = 160,527.

Figure 1.H.5: Do firms respond to retirement delays happening in the future?

(a) Layoffs



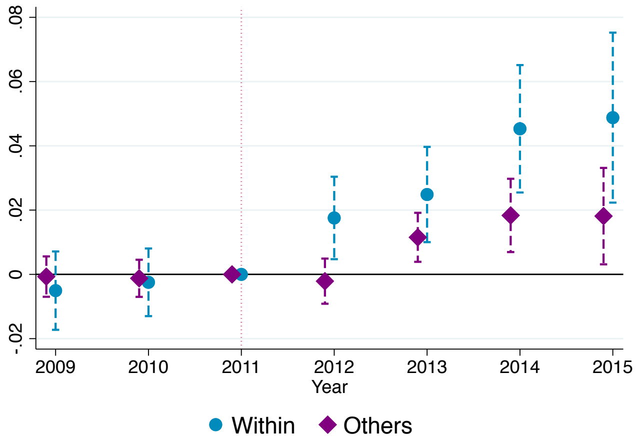
(b) New hires



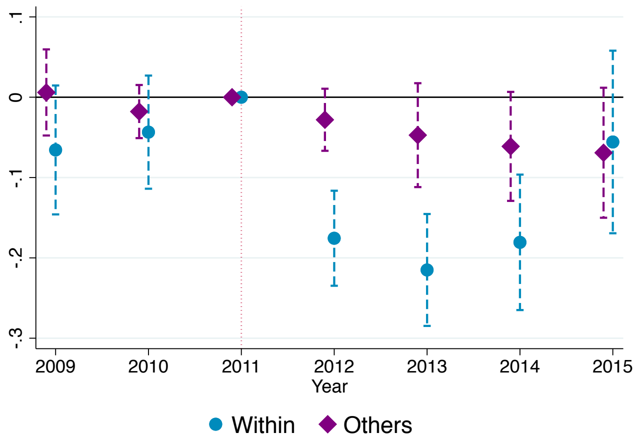
Notes: The figure plots results of a modified version of specification (1.3). Whiskers represent 95 per cent confidence intervals. Standard errors are clustered at the firm level. The treatment is the firm-level average change in the full retirement date of full-time workers who were expected to retire either in 2015 or in 2016. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; (iii) employed at least one worker who was expected to retire in 2015 or 2016; and (iv) employed no *potential retirees*, *i.e.* no full-time workers who were expected to retire by 2014. Number of observations = 265,216. Mean outcomes pre-reform: layoffs = 0.36; new hires = 3.5.

Figure 1.H.6: Labour demand adjustments within and across occupation groups

(a) Layoffs



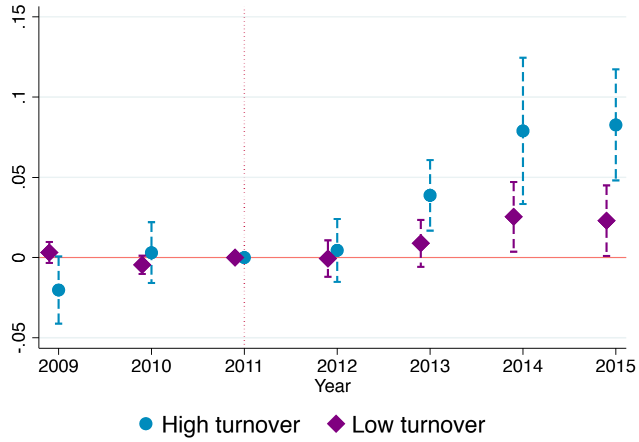
(b) New hires



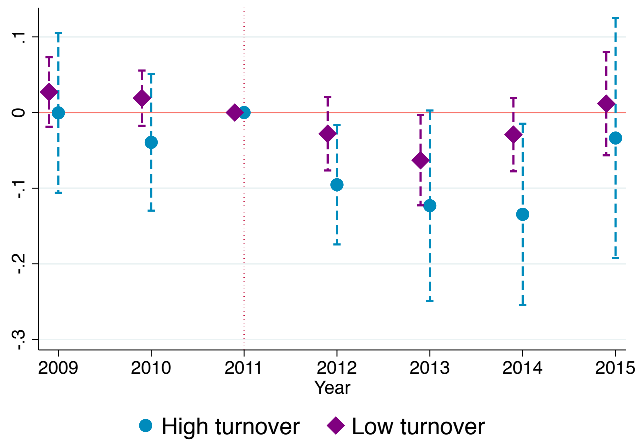
Notes: The figure plots the results of a modified version of specification (1.3), where the unit of analysis is the firm-occupation group and there are two treatments. The specification also features occupation-year fixed effects and standard errors are clustered at the firm-occupation level. Whiskers represent 95 per cent confidence intervals. The occupation groups are blue-collars, white-collars and managers. The first treatment is the average change in the full retirement date of *potential retirees* employed at the firm in the same occupation group when the reform is passed ("Within"). The second treatment is the average change of the full retirement date of *potential retirees* who work in any other occupation group within the firm ("Others"). We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; (iii) had at least two occupation groups with at least three workers in 2009; and (iv) employed at least one *potential retiree* in the last quarter of 2011. Number of observations = 580,734. Mean outcomes pre-reform at the firm-occupation level: layoffs = 0.14; new hires = 2.18.

Figure 1.H.7: Labour demand adjustments in firms with high and low turnover rates in the pre-reform period

(a) Layoffs



(b) New hires



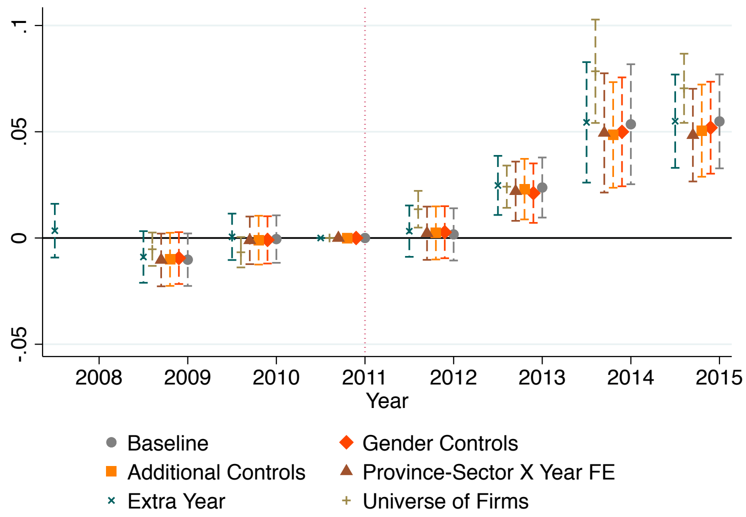
Notes: The figure shows the heterogeneous effect of a 1-year increase of T_i on firms with high (above the median) and low (below the median) turnover rates in the pre-reform period. Whiskers represent 95 per cent confidence intervals. Standard errors are clustered at the firm level. Estimates are based on a modified version of specification (1.3), which reads:

$$Y_{it} = \lambda_i + \gamma_t + \sum_{k=2009}^{2015} \beta_k^T \mathbb{1}(k=t) \times T_i + \sum_{k=2009}^{2015} \beta_k^{T,to} \mathbb{1}(k=t) \times T_i \times TO_i + \sum_{k=2009}^{2015} \beta_k^{to} \mathbb{1}(k=t) \times TO_i + \varepsilon_{it}$$

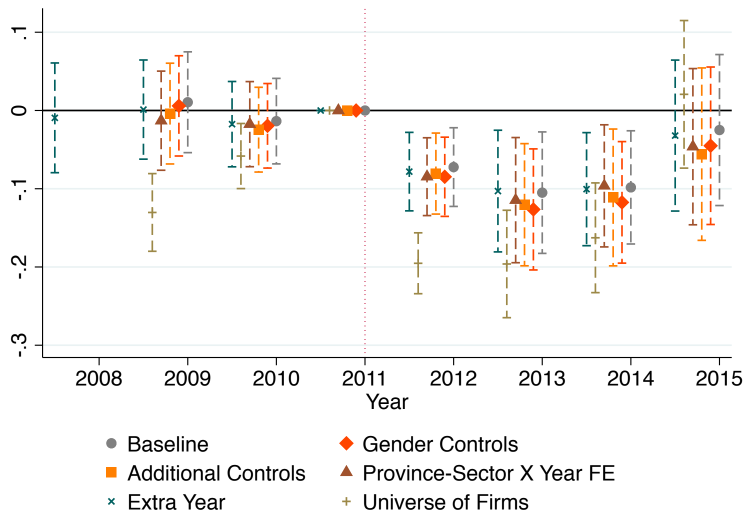
Purple dots plot the coefficients β_k^Ts, while light blue dots plot the linear combinations $\beta_k^T + \beta_k^{T,to}$. TO_i is a dummy that takes value 1 if firm i belongs to the top half of the distribution of pre-refrom turnover rates. We define the turnover rate as the average number of layoffs, quits and terminations of temporary contracts in the pre-reform period (2009-2011), normalized by employment at the beginning of the period (first quarter of 2009). T_i is the average change in the full retirement date of *potential retirees* employed at a given firm when the pension reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within three years (by 2014) under pre-reform rules when the reform is implemented. The regression is estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. Number of observations = 430,038. Mean outcome pre-reform: layoffs = 0.39; new hires = 4.79.

Figure 1.H.8: Labour demand adjustments: robustness to alternative specifications and samples

(a) Layoffs

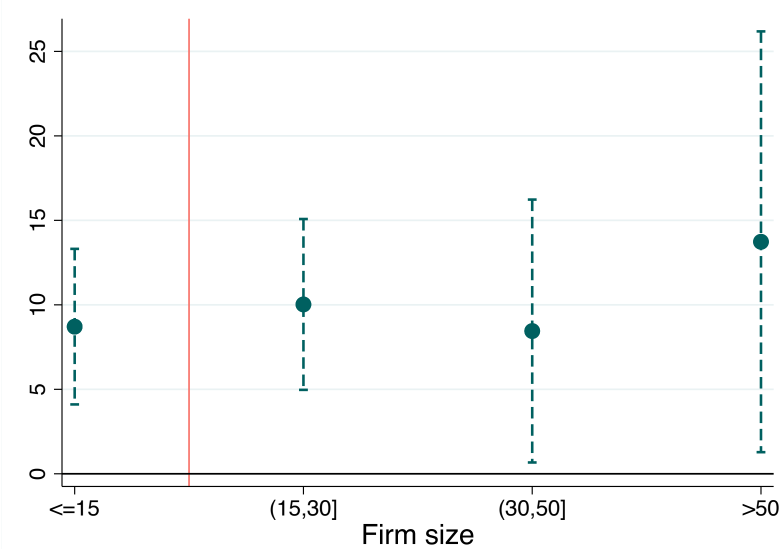


(b) New hires



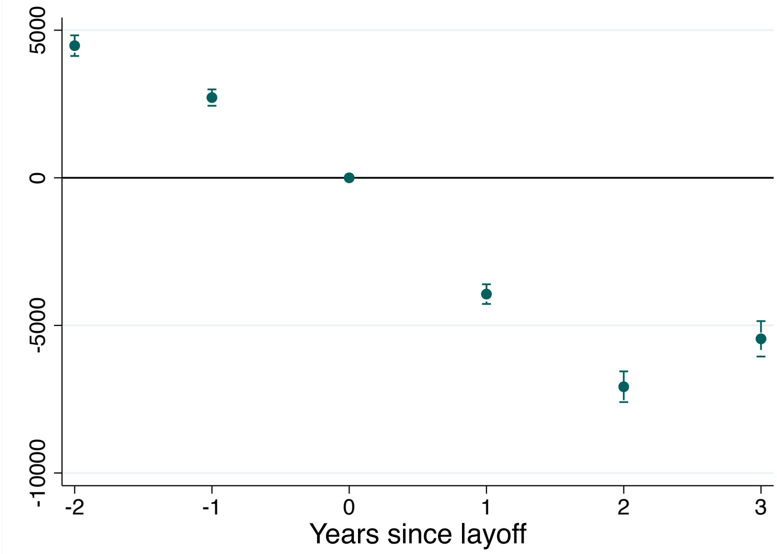
Notes: The figure shows the robustness of the main estimates on layoffs (Panel a) and new hires (Panel b). Whiskers represent 95 per cent confidence intervals. Standard errors are clustered at the firm level. Estimates based on specification (1.3) and reported in figures 1.3 and 1.4 ("Baseline") are confronted with estimates from augmented versions of specification (1.3) that feature: (i) the interaction between year fixed effects and quintiles of the share of female workers at the firm ("Gender controls"); (ii) the interaction between year fixed effects and quintiles of the share of female workers, firm size, firm age, firm average wage, the share of young (≤ 35), middle-aged ($36-55$) and older (> 55) workers ("Additional controls"); (iii) province-year fixed effects and sector-year fixed effects ("Province-Sector $\times$ Year FE"). Furthermore, we estimate versions of specification (1.3) that: (iv) include year 2008 ("Extra Year"); (v) include firms that did not employ any *potential retiree* at the time of the reform ("Universe of Firms").

Figure 1.H.9: Labour demand adjustments: layoffs by firm size



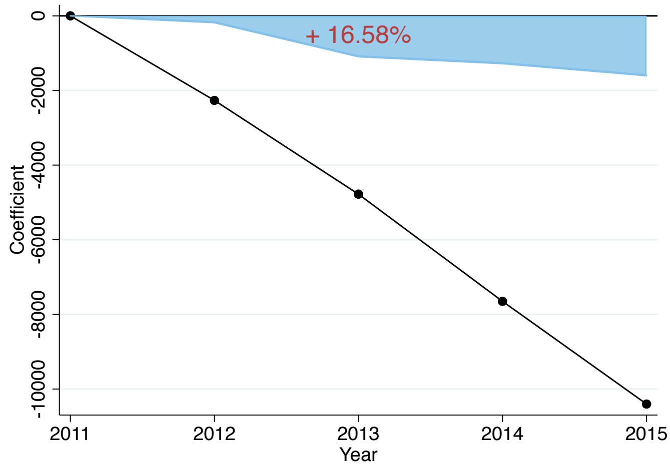
Notes: The figure shows the post-reform percentage change in layoffs by firm size at the beginning of the period (first quarter of 2009). The plotted coefficients are obtained by re-scaling the coefficients associated to the term $T_i \times Post_t$ in specification (1.4) by the pre-reform mean outcome. Whiskers represent 95 per cent confidence intervals. Standard errors are clustered at the firm level. The vertical red line separates firms where a more stringent dismissal discipline applies (> 15 employees) from other firms. The treatment T_i is defined as the average change in the full retirement date of *potential retirees* employed at the firm when the pension reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within three years (by 2014) under pre-reform rules when the reform is implemented. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first period of 2009; and (iii) employed at least one *potential retiree* in the last period of 2011. Observations: [3,15] = 239,995; (15,30] = 81,354; (30,50] = 50,295; (50,200] = 58,394. Mean outcome pre-reform: [3,15] = 0.26; (15,30] = 0.39; (30,50] = 0.51; (50,200] = 0.78.

Figure 1.H.10: Cost of Layoffs



Notes: The figure shows the effect of separating from a firm because of a layoff on subsequent labour earnings. Whiskers represent 95 per cent confidence intervals. Coefficients come from a difference-in-differences specification estimated on a sample constructed by matching fired individuals to similar workers who did not separate from the firm where they were incumbent at the time of the reform. The matching is achieved through a coarsened exact matching procedure. The matching covariates are: age, sex, wage, occupation, a dummy for permanent contracts, a dummy for full-time contracts, experience, sector, province and firm size. Mean outcome pre-reform = 28,532.71.

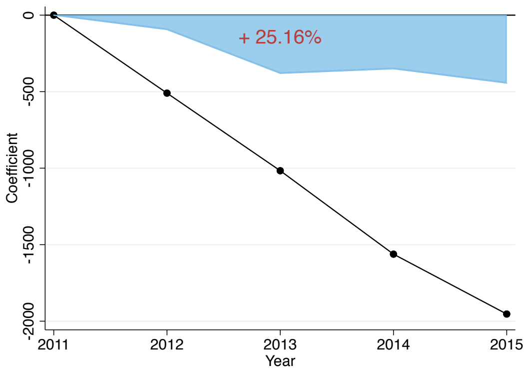
Figure 1.H.11: Co-workers' earnings: the cost of layoffs vs within-firm dynamics



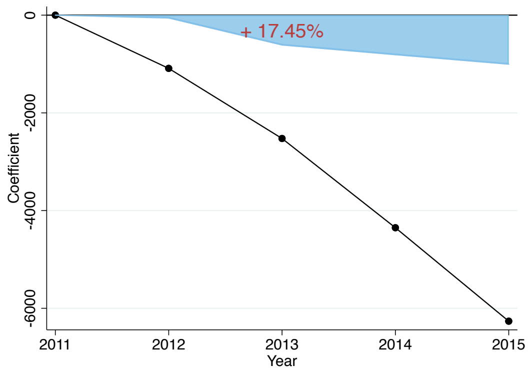
Notes: The figure shows the share of the earnings decline experienced by full-time permanent co-workers that can be imputed to layoffs. Co-workers are full-time permanent employees who were not expected to retire within 3 years when the reform is passed and who worked in a firm with at least one *potential retiree* belonging to the sample described in Subsection 1.5.2. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Worker-level observations are aggregated at the level of the firm where workers were incumbents in the last quarter of 2011. The black line represents the effect of a 1-year increase of T_i on total labour earnings. Labor earnings are winsorized at the 99th percentile. The blue-shaded area represents the share of earnings loss imputed to layoffs. This area is the result of computations combining estimates of the cost of job losses and estimates of the effect of the reform on total layoffs (Section 1.9.1).

Figure 1.H.12: Co-workers' earnings by age: the cost of layoffs vs within-firm dynamics

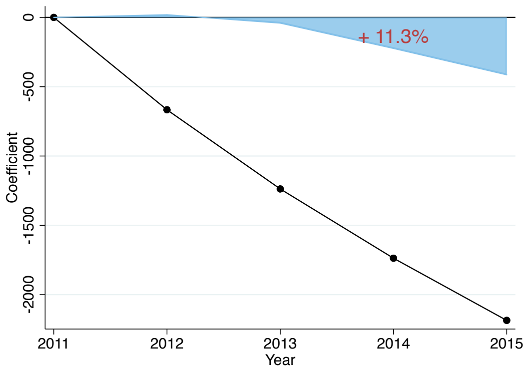
(a) Young



(b) Middle-aged



(c) Old



Notes: The figure shows the share of the earnings decline experienced by full-time permanent co-workers that can be imputed to layoffs for young workers (≤ 35, Panel a), middle-aged workers (aged 36-55, Panel b), and older workers (aged over 55, Panel c). See notes to Figure 1.H.11 for more details.

Table 1.H.1: Pre- and post-reform pension requirements - Additional Details

	Panel (a): Old-age pension			
	Men		Women	
	Pre-reform	Post-reform	Pre-reform	Post-reform
	Age requirement			
2011	65YA	Not in place	60YA	Not in place
2012	65YA	66YA	60YA	62YA
2013	65YA+3MA	66YA+3MA	60YA+3MA	62YA+3MA
2014	65YA+3MA	66YA+3MA	60YA+4MA	63YA+9MA
2015	65YA+3MA	66YA+3MA	60YA+6MA	63YA+9MA
2016	65YA+7MA	66YA+7MA	61YA+1MA	65YA+7MA
2017	65YA+7MA	66YA+7MA	61YA+5MA	65YA+7MA
2018	65YA+7MA	66YA+7MA	61YA+10MA	66YA+7MA
	Contribution requirement			
	20YC	20YC	20YC	20YC
	Waiting window			
	12 months	No	12 months	No

	Panel (b): Seniority pension			
	Pre-reform		Post-reform	
			Men	Women
2011	Quota 96 (60YA and 35 YC)	or 40 YC	Not in place	
2012	Quota 96 (60YA and 35 YC)	or 40 YC	42YC+1MC	41YC+1MC
2013	Quota 97.3 (61YA+3MA and 35 YC)	or 40 YC	42YC+5MC	41YC+5MC
2014	Quota 97.3 (61YA+3MA and 35 YC)	or 40 YC	42YC+6MC	41YC+6MC
2015	Quota 97.3 (61YA+3MA and 35 YC)	or 40 YC	42YC+6MC	41YC+6MC
2016	Quota 97.6 (61YA+7MA and 35 YC)	or 40 YC	42YC+10MC	41YC+10MC
2017	Quota 97.6 (61YA+7MA and 35 YC)	or 40 YC	42YC+10MC	41YC+10MC
2018	Quota 97.6 (61YA+7MA and 35 YC)	or 40 YC	42YC+10MC	41YC+10MC
	Waiting window			
	12 months		No	

Notes: The table reports the requirements for claiming old-age (Panel a) and seniority (Panel b) pensions under pre-reform rules - had they remained in place - and under post-reform rules, for private sector employees over the period 2012-2018. It takes into account the anticipated upward adjustments due to increased life expectancy that took place in 2013 and 2016. YA and MA flag the age requirement in terms of years and months, respectively. YC and MC flag the contribution requirement in terms of years and months, respectively. The "waiting window" is the period that elapses between the date when an individual becomes eligible to claim a pension and the date when he/she can collect the first pension benefits. The "waiting window" was set to 12 months in 2011 for old-age pensions and seniority pensions under the "quota" system. For those retiring upon reaching 40 years of contributions, the "waiting window" was 13 months in 2012, 14 months in 2013 and 15 months from 2014 onward. The "waiting window" has been abolished by the *Fornero* reform. For workers with at least 15 years of contributions by 1992, the contribution requirement of the old-age pension is 15 years rather than 20. Very few workers, among those who were old enough to claim an old-age pension by 2014 under pre-reform rules and had at least 15 years of contributions in 1992, had accrued less than 20 years of contribution by 2011. We therefore do not take into account this exception.

Table 1.H.2: Placebo tests

	Layoffs (1)	New Hires (2)
Panel (a)		
$T_i \cdot$ Post 2009	0.013*	-0.023
	(0.007)	(0.036)
$(R_i/N_i) \cdot$ Post 2009	0.002	-0.315***
	(0.005)	(0.081)
Panel (b)		
$T_i \cdot$ Post 2010	0.007	0.002
	(0.007)	(0.037)
$(R_i/N_i) \cdot$ Post 2010	-0.010**	-0.277***
	(0.005)	(0.060)
Observations	184,302	184,302
Mean outcome pre-2012	0.39	4.79
Mean T_i	1.37	1.37
Std. Dev. T_i	1.33	1.33
Mean R_i/N_i	0.06	0.06
Std. Dev. R_i/N_i	0.09	0.09

Notes: The table reports the coefficients from a set of placebo tests whereby we re-allocate the reform date to December 2009 (Panel a) or December 2010 (Panel b). The first row of every panel reports the effect of a 1σ increase of T_i, which is the average change in the full retirement date of *potential retirees* employed at the firm when the reform is passed. The second row of every panel reports the effect of a 1σ increase of R_i/N_i, which is the number of *potential retirees* who delay retirement by at least one year with respect to the pre-reform predicted full retirement date, divided by employment. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Estimates come from specification (1.4), restricting the sample to the pre-reform period (2009-2011). The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. *, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.3: Balancing tests

	T_i (1)	R_i/N_i (2)	Mean (3)
Share male employees	-0.011***	0.004***	0.66
	(0.001)	(0.001)	
Share employees aged $\leq$ 35	0.000	-0.017***	0.30
	(0.001)	(0.001)	
Share employees aged (35,55]	-0.011***	-0.011***	0.58
	(0.001)	(0.001)	
Share employees aged $>$ 55	0.010***	0.028***	0.12
	(0.001)	(0.001)	
Employees tenure	-0.105***	0.295***	8.21
	(0.020)	(0.024)	
Employees experience	-0.132***	0.324***	16.48
	(0.002)	(0.002)	
Employment	0.492***	-7.237***	25.80
	(0.114)	(0.152)	
Average daily wage	0.665	-1.574***	92.05
	(0.614)	(0.373)	
Share temporary employees	0.003***	-0.006***	0.08
	(0.001)	(0.001)	
Share full-time employees	-0.002*	-0.006***	0.89
	(0.001)	(0.001)	
Share blue-collars	-0.002	0.010***	0.61
	(0.001)	(0.001)	
Share white-collars	0.001	-0.009***	0.33
	(0.001)	(0.001)	
Share managers	0.001	-0.002***	0.02
	(0.000)	(0.000)	
Firm age	-0.151**	-0.333***	19.96
	(0.051)	(0.054)	
Observations	60195	60195	

Notes: The table reports a set of balancing tests whereby firms' characteristics at the beginning of the period (first quarter of 2009) are regressed on: (i) T_i, the change in the full retirement date of *potential retirees* employed at the firm when the reform is passed (column 1); (ii) R_i/N_i, the number of *potential retirees* who delay retirement by at least one year with respect to the predicted pre-reform full retirement date, divided by employment (column 2). Both T_i and R_i/N_i are standardized to have mean equal to 0 and standard deviation equal to 1. We define *potential retirees* as those full-time workers who were expected to retire within three years (by 2014) under pre-reform rules when the reform is implemented. All regressions feature sector-province fixed effects, which in the cross-section of firms leads to the drop of some singleton observations from the estimation. Standard errors reported in parenthesis are clustered at the sector-province level. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.4: Firms with at least one *potential retiree* and other firms

	Sample		Other Firms	
	mean	std. dev.	mean	std. dev.
	(1)	(2)	(3)	(4)
Firm size	25.85	32.05	8.06	9.83
Firm age	19.93	12.78	14.07	10.68
Share in manufacturing	0.44	0.50	0.26	0.44
Share in services	0.34	0.47	0.51	0.50
Share male employees	0.66	0.30	0.55	0.35
Share employees aged ≤ 35	0.30	0.19	0.46	0.28
Share employees aged (35-55]	0.58	0.19	0.49	0.27
Share employees aged > 55	0.12	0.12	0.05	0.11
Avg. employees tenure	8.21	4.82	5.67	4.13
Avg. employees experience	16.46	4.21	12.41	4.94
Share blue collars	0.61	0.32	0.56	0.37
Share white collars	0.33	0.30	0.34	0.36
Share managers	0.02	0.07	0.01	0.06
Share full-time contracts	0.89	0.17	0.74	0.30
Share permanent contracts	0.92	0.15	0.90	0.19
Avg. real daily wage	92.04	146.07	78.89	188.87
Observations	61,434		333,800	

Notes: The table reports descriptive statistics for firms in our sample, as well as for other firms with a single Contribution Identification Number, in the same size class (3-200) and that remain active throughout the period 2009-2015. Firms' characteristics are measured at the beginning of the period (first quarter of 2009). Workers' tenure and experience are truncated at 27 years, because matched employer-employee data have been available since 1983. Firms in our sample have a single Contribution Identification Number and: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.

Table 1.H.5: *Potential retirees* and other workers

	Potential retirees		Other workers	
	mean	std. dev.	mean	std. dev.
	(1)	(2)	(3)	(4)
Male	0.71	0.46	0.71	0.45
Age	57.77	2.89	40.91	9.77
Tenure	14.24	9.37	8.81	7.44
Experience in private sector	23.86	8.60	15.30	9.81
Blue collar	0.66	0.47	0.60	0.49
White collar	0.29	0.45	0.34	0.47
Manager	0.05	0.21	0.04	0.19
Open-ended contract	0.96	0.20	0.90	0.30
Daily gross real wage	109.74	114.52	102.17	111.78
Observations	98,358		1,434,381	

Notes: The table reports the characteristics at the beginning of the period (first quarter of 2009) of *potential retirees* and full-time co-workers employed in firms belonging to our sample. Tenure and experience are truncated at 27 years, because matched employer-employee data have been available since 1983 only. Firms in our sample have a single Contribution Identification Number and: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.

Table 1.H.6: *Potential retirees* and other similar older workers: absences from work

	Other older workers		*Potential retirees*		Difference
	Mean	Std. Dev.	Mean	Std. Dev.	
	(1)	(2)	(3)	(4)	(5)
Prob. sickness	0.30	0.46	0.35	0.40	0.05***
Prob. work-related injury	0.06	0.23	0.07	0.25	0.01***
Prob. leave	0.03	0.18	0.04	0.21	0.01***
Observations	550,108		93,835		

Notes: The table reports the probability of being absent from work due to sickness, work-related injuries or leave in the pre-reform period (2009-2011), for *potential retirees* and similar older employees (aged 50 and above) who are matched - via an exact matching procedure - along several dimensions. Matching covariates are: age, experience, gender, type of contract (permanent vs. temporary), occupation (blue-collar, white-collar, manager), as well as firm's province, sector and size. We find a match for 95.4 per cent of all *potential retirees* in our sample. The last column reports the difference between the third and the first column.
*, **, *** indicate that the difference is significant at the 10, 5 and 1 per cent level, respectively.

Table 1.H.7: The effect of the reform on the working lives of *potential retirees*

	Months Worked	Months in Retirement	Months in Retirement (Benchmark)
	(1)	(2)	(3)
$\delta_\theta \cdot$ Post	1.250***	-1.787***	-2.136***
	(0.011)	(0.009)	(0.009)
Observations	688,492	688,492	688,492
Mean outcome pre-2012	11.28	0	0
Mean δ_θ	1.36	1.36	1.36
Std. Dev. δ_θ	1.40	1.40	1.40

Notes: The table reports estimates from a richer individual-level version of specification (1.4), where the unit of analysis is the single *potential retiree* in our sample of firms, the treatment is δ_θ and there are individual, age-year and gender-year fixed effects. Standard errors are clustered at the individual level. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Coefficients measure the effect of a 1-year shift in the full retirement date δ_θ (as defined in equation (1.1)) on months worked, months in retirement, and predicted (benchmark) months in retirement if the workers retired at the post-reform predicted date. We define an individual as retired when he/she starts collecting retirement benefits.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.8: Labour demand adjustments: layoffs

							Layoffs probability	
	All	Young	Middle-aged	Old	Old Non-*potential retirees*	Young	Middle-aged	Old
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$T_i \cdot \mathbb{1}(\text{Year} = 2009)$	-0.0102	-0.0052*	-0.0067*	0.0017	-0.0009	-0.0027*	-0.0014*	0.0010
	(0.0063)	(0.0028)	(0.0038)	(0.0013)	(0.0010)	(0.0015)	(0.0008)	(0.0009)
$T_i \cdot \mathbb{1}(\text{Year} = 2010)$	-0.0005	-0.0031	-0.0004	0.0030**	0.0007	-0.0015	-0.0005	0.0017**
	(0.0057)	(0.0023)	(0.0037)	(0.0013)	(0.0010)	(0.0013)	(0.0008)	(0.0008)
$T_i \cdot \mathbb{1}(\text{Year} = 2012)$	0.0017	0.0006	0.0027	-0.0017	-0.0007	-0.0003	-0.0001	-0.0015
	(0.0063)	(0.0025)	(0.0040)	(0.0016)	(0.0012)	(0.0009)	(0.0009)	(0.0013)
$T_i \cdot \mathbb{1}(\text{Year} = 2013)$	0.0238***	0.0067**	0.0096**	0.0075***	0.0027**	0.0015	0.0001	0.0046***
	(0.0072)	(0.0027)	(0.0045)	(0.0016)	(0.0012)	(0.0011)	(0.0009)	(0.0012)
$T_i \cdot \mathbb{1}(\text{Year} = 2014)$	0.0535***	0.0127***	0.0262***	0.0146***	0.0071***	0.0053**	0.0056	0.0082***
	(0.0144)	(0.0036)	(0.0096)	(0.0025)	(0.0021)	(0.0022)	(0.0041)	(0.0017)
$T_i \cdot \mathbb{1}(\text{Year} = 2015)$	0.0549***	0.0094***	0.0256***	0.0198***	0.0084***	0.0034**	0.0034***	0.0106***
	(0.0113)	(0.0030)	(0.0073)	(0.0026)	(0.0022)	(0.0016)	(0.0011)	(0.0016)
Observations	430,038	430,038	430,038	430,038	430,038	305,319	305,319	305,319
Mean outcome pre-2012	0.39	0.13	0.2	0.06	0.04	0.15	0.24	0.07
Mean T_i	1.37	1.37	1.37	1.37	1.37	1.44	1.44	1.44
Std. Dev. T_i	1.33	1.33	1.33	1.33	1.33	1.38	1.38	1.38

Notes: The table reports estimates from specification (1.3), which includes firm and year fixed effects. Standard errors in parentheses are clustered at the firm level. Coefficients $T_i \cdot \mathbb{1}(\text{Year} = k)$ show how a 1-year increase in T_i affects layoffs in year k compared to year 2011, when the effect is normalized to 0. Column 1 shows the effect on total layoffs. Columns 2 to 5 report the effect on layoffs of young (35 or below), middle-aged (36-55), old (over 55) and old non-*potential retirees*, respectively. Columns 6 to 8 report the effect on the layoff probability of young, middle-aged and old workers, respectively. The layoff probability is defined as the ratio between the number of layoffs in every age group in a given year and the respective number of incumbent employees at the beginning of the period. The treatment T_i is defined as the average change in the full retirement date of *potential retirees* employed at the firm when the reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.9: Labour demand adjustments: new hires

	All (1)	Young (2)	Middle-aged (3)	Old (4)	Permanent (5)	Temporary (6)
$T_i \cdot \mathbb{1}(\text{Year} = 2009)$	0.011	0.013	-0.002	-0.001	0.003	0.008
	(0.033)	(0.017)	(0.017)	(0.004)	(0.018)	(0.025)
$T_i \cdot \mathbb{1}(\text{Year} = 2010)$	-0.014	-0.003	-0.009	-0.002	0.003	-0.017
	(0.028)	(0.014)	(0.015)	(0.004)	(0.015)	(0.022)
$T_i \cdot \mathbb{1}(\text{Year} = 2012)$	-0.072***	-0.030*	-0.033**	-0.010***	-0.008	-0.065***
	(0.026)	(0.015)	(0.013)	(0.004)	(0.014)	(0.022)
$T_i \cdot \mathbb{1}(\text{Year} = 2013)$	-0.105***	-0.055***	-0.045**	-0.005	-0.027	-0.078**
	(0.040)	(0.020)	(0.022)	(0.005)	(0.022)	(0.031)
$T_i \cdot \mathbb{1}(\text{Year} = 2014)$	-0.098***	-0.049**	-0.049***	-0.000	-0.012	-0.086***
	(0.037)	(0.021)	(0.017)	(0.005)	(0.015)	(0.032)
$T_i \cdot \mathbb{1}(\text{Year} = 2015)$	-0.025	-0.021	-0.014	0.010	0.046*	-0.071*
	(0.049)	(0.026)	(0.024)	(0.007)	(0.024)	(0.039)
Observations	430,038	430,038	430,038	430,038	430,038	430,038
Mean Outcome pre-2012	4.79	2.38	2.06	0.35	1.46	3.32
Mean T_i	1.37	1.37	1.37	1.37	1.37	1.37
Std. Dev. T_i	1.33	1.33	1.33	1.33	1.33	1.33

Notes: The table reports estimates from specification (1.3), which includes firm and year fixed effects. Standard errors in parentheses are clustered at the firm level. Coefficients $T_i \cdot \mathbb{1}(\text{Year} = k)$ show how a 1-year increase in T_i affects hiring in year k compared to year 2011, when the effect is normalized to 0. Column 1 shows the effect on total new hires. Columns 2 to 4 report the effects on new hires of young (35 or below), middle-aged (36-55) and old (over 55) workers, respectively. Columns 5 and 6 report the effect of new hires under permanent and temporary contracts, respectively. The treatment T_i is defined as the average change in the full retirement date of *potential retirees* employed at the firm when the reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.10: Do firms respond to retirement delays happening in the future?

	Layoffs (1)	New Hires (2)
$T_i^{45} \cdot \text{Post}$	0.0076	-0.030
	(0.0050)	(0.028)
Observations	265,216	265,216
Mean outcome pre-2012	0.36	3.5
Mean T_i^{45}	2.01	2.01
Std. Dev. T_i^{45}	1.45	1.45

Notes: The table reports the results of a modified version of specification (1.4). Standard errors in parentheses are clustered at the firm level. The treatment is the firm-level average change in the full retirement date of full-time workers who were expected to retire in either 2015 or 2016 when the reform is passed (T_i^{45}). The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; (iii) employed at least one worker who was expected to retire in 2015 or 2016 in the last quarter of 2011; and (iv) employed no *potential retiree*, *i.e.* no full-time worker who was expected to retire by 2014.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.11: Labour demand adjustments within and across occupation groups

	Layoffs (1)	New Hires (2)
$T_i^{\text{within}} \cdot \text{Post}$	0.037***	-0.120***
	(0.007)	(0.040)
$T_i^{\text{others}} \cdot \text{Post}$	0.012***	-0.047
	(0.003)	(0.030)
Observations	580,734	580,734
Mean outcome pre-2012	0.14	2.18
Mean T_i^{within}	0.57	0.57
Std. Dev. T_i^{within}	1.04	1.04
Mean T_i^{others}	1.02	1.02
Std. Dev. T_i^{others}	1.2	1.2
P-value one-sided test	0.00	0.07

Notes: The table reports the results of a modified version of specification (1.4), where the unit of analysis is the firm-occupation group and there are two treatments. The specification also features occupation-year fixed effects and standard errors are clustered at the firm-occupation level. The occupation groups are blue-collars, white-collars and managers. The first treatment is the average change of the full retirement date of *potential retirees* employed at the firm in the same occupation group when the reform is passed (T_i^{within}). The second treatment is the average change of the full retirement date of *potential retirees* who work in any other occupation group within the firm (T_i^{others}). We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. The table also reports the p-value of a one-sided test for the difference between the two coefficients. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; (iii) had at least two occupation groups with at least three workers in 2009; and (iv) employed at least one *potential retiree* in the last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.12: Labour demand adjustments in firms with high and low turnover rates in the pre-reform period

	Layoffs (1)	New hires (2)
$T_i \cdot$ Post	0.015***	-0.043**
	(0.005)	(0.022)
$T_i \cdot$ Post $\cdot$ High Turnover	0.042***	-0.041
	(0.012)	(0.060)
Observations	430,038	430,038
Mean outcome pre-2012	0.39	4.79
Mean T_i	1.37	1.37
Std. Dev. T_i	1.33	1.33

Notes: The table reports the heterogeneous effect of a 1-year increase of T_i on layoffs and new hires in firms with high (above the median) and low (below the median) turnover rates in the pre-reform period. Estimates come from a modified version of specification (1.4), which reads:

$$Y_{it} = \lambda_i + \gamma_t + \beta^T T_i \cdot Post_t + \beta^{T,to} T_i \cdot Post_t \times TO_i + \beta^{to} Post_t \cdot TO_i + \varepsilon_{it}$$

Standard errors are clustered at the firm level. T_i is the average change in the full retirement date of *potential retirees* employed at a given firm when the pension reform is passed. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. TO_i is a dummy variable that takes value 1 if firm i belongs to the top half of the distribution of turnover in the pre-reform period. Turnover is defined as the average share of layoffs, quits and terminations of temporary contracts in the pre-reform period (2009-2011), normalized by employment at the beginning of the period (first quarter of 2009). The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.13: Incumbents' labour earnings

	Co-workers					*Potential retirees*
	All	All (w/ non-work subsidies)	Permanent employees Young	Middle-aged	Old	All
	(1)	(2)	(3)	(4)	(5)	(6)
$T_i \cdot \mathbb{1}(\text{Year} = 2009)$	-428.6	-469.9	-164.1	1,777.9	102.1	180.0**
	(1909.0)	(1903.4)	(659.1)	(1210.6)	(169.6)	(75.7)
$T_i \cdot \mathbb{1}(\text{Year} = 2010)$	-181.9	-205.5	42.9	1,252.8**	90.9	28.4
	(968.1)	(952.16)	(334.2)	(563.7)	(121.1)	(54.0)
$T_i \cdot \mathbb{1}(\text{Year} = 2012)$	-2,262.7**	-1,822.5**	-509.5**	-1,089.1*	-666.4***	2050.0***
	(889.5)	(857.6)	(235.2)	(599.7)	(192.4)	(110.5)
$T_i \cdot \mathbb{1}(\text{Year} = 2013)$	-5,220.0***	-3,824.3***	-1,016.7***	-2,524.4**	-1,237.2***	5160.2***
	(1605.0)	(1367.7)	(383.9)	(1117.2)	(313.4)	(225.9)
$T_i \cdot \mathbb{1}(\text{Year} = 2014)$	-8,578.1***	-6,346.5***	-1,561.9***	-4,350.8***	-1,736.9***	6296.1***
	(1990.8)	(1737.6)	(518.3)	(1355.2)	(395.4)	(342.0)
$T_i \cdot \mathbb{1}(\text{Year} = 2015)$	-11,356.0***	-8,515.2***	-1,953.2***	-6,263.7***	-2,185.7***	7202.1***
	(2359.5)	(2089.2)	(623.8)	(1595.7)	(499.8)	(435.6)
Observations	423,346	423,346	419,363	419,363	419,363	425,971
Mean outcome pre-2012	647,005.96	648,244.10	138,497.38	440,187.65	42,319.77	49,951.32
Mean T_i	1.36	1.36	1.36	1.36	1.36	1.36
Std. Dev. T_i	1.32	1.32	1.31	1.31	1.31	1.33

Notes: The table shows the response of labour earnings of *potential retirees* and of their full-time co-workers to a 1-year increase in T_i, the shift in the full retirement date of *potential retirees* employed at the firm when the reform is passed. Standard errors in parentheses are clustered at the firm level. Worker-level observations are aggregated at the level of the firm where workers were incumbents in the last quarter of 2011. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Column 1 and 3-6 do not include non-work subsidies, while column 2 does. Earnings and non-work subsidies are winsorized at the 99th and the 99.99th percentiles, respectively. Columns 3 to 5 display the effect on permanent young ($\leq$ 35), middle-aged (36-55), and older (over 55) co-workers, respectively. The coefficients $T_i \cdot \mathbb{1}(\text{Year} = k)$ show the effect of a 1-year increase in T_i on labour earnings in year k relative to year 2011, when the effect is normalized to zero. The regressions are based on specification (1.3) and are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011. Observations are weighted according to firm size at the beginning of the period.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.14: The effect of the reform on *potential retirees* and co-workers

	Potential retirees			Co-workers		
	Labour Earnings (1)	Total Transfers (2)	Pension Payments (3)	Labour Earnings (4)	Total Transfers (5)	Pension Payments (6)
$T_i \cdot$ Post	5,107.6***	315.8***	-4,685.9***	-6,650.7***	1,775.8***	-75.8
	(251.5)	(37.4)	(132.4)	(1,787.5)	(534.0)	(161.1)
Observations	425,971	425,971	425,971	423,346	423,346	423,346
Mean outcome pre-2012	49,951.32	303.21	4.2	647,005.96	1,702.16	1,337.52

Notes: The table reports the response of labour earnings, pension benefits and total social insurance transfers (non-work subsidies and disability benefits) to a 1-year increase of T_i, the shift in the full retirement date of *potential retirees* employed at the firm when the reform is passed. Standard errors are clustered at the firm level. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented. Worker-level observations are aggregated at the level of the firm where workers were incumbents in the last quarter of 2011. Columns 1 to 3 show the effect on *potential retirees*, columns 4 to 6 on their co-workers. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in the last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 1.H.15: The components of the fiscal externality

	Potential retirees (1)	Co-workers (2)
Labour earnings	20,708.28***	-27,416.8***
	(1,076.507)	(6,281.804)
Pension entitlements	-18,752.2***	81.57
	(528.340)	(564.589)
Disability benefits	198.010***	112.562
	(22.150)	(158.488)
Non-work Subsidies	1,054.629***	6,908.183***
	(152.693)	(2,093.375)
Early retirement (months)	2.15***	0.713***
	(0.000)	(0.270)

Notes: The table reports the sum of coefficients $\{\beta_k^T\}_{k=2012}^{k=2015}$ from specification (1.3). The treatment T_i is defined as the average shift in the full retirement date of *potential retirees* employed at the firm when the reform is implemented. Standard errors in parentheses are clustered at the firm level. We define *potential retirees* as those full-time workers who were expected to retire within 3 years (by 2014) under pre-reform rules when the reform is implemented.Worker-level observations are aggregated at the level of the firm where workers were incumbents in the last quarter of 2011. Column 1 reports the estimates for the sample of *potential retirees*, while column 2 displays the effect on their co-workers. Labour earnings, pension benefits and disability benefits are winsorized at the 99th percentile, while non-work subsidies are winsorized at the 99.99th percentile. The regressions are estimated on the universe of private sector firms with a single Contribution Identification Number that: (i) were active every year in the period 2009-2015; (ii) employed between 3 and 200 employees in the first quarter of 2009; and (iii) employed at least one *potential retiree* in last quarter of 2011.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Chapter 2

Pension reforms and labour supply: evidence from Italy

2.1 Introduction

Population ageing is one of the main challenges that many countries are confronted with. The old-age dependency ratio in OECD countries was equal to 27.9 in 2015, up from 19.5 in 1975, and is projected to double by 2075 (OECD, 2017).[1] This demographic transformation exerts mounting pressure on the financial sustainability of public pension systems: over the 1980-2015 period, public pension spending has increased from 5.5 per cent to 7.5 per cent of GDP. Spurred by these trends, many countries have implemented or are considering enacting pension reforms that aim to encourage elderly labour force participation and contain pension expenditures, by reducing the generosity of retirement benefits or tightening eligibility criteria.

A rich literature has explored how these changes in the public pension systems affect employment at older ages, as well as the retirement and benefit claiming decisions. Individuals, however, sometimes become aware of modifications of retirement rules at earlier stages of their life and before they come near to making those decisions. A feature common to many reforms is, indeed, the presence of grandfathering clauses and gradual phase-in schedules; this means that new provisions do not apply to workers on the cusp of retirement, while they partially or fully affect younger ones.[2] In such circumstances, do individuals start changing their behaviour in the aftermath of the reform announcement? In particular, do they adjust their labour supply and how do adjustments evolve over time? Have, therefore, pension reforms broader and further labour supply effects than those documented at the very end of an individual's active life? These are the questions that I attempt to address in this paper.

To do so, I study the labour supply dynamics of a sample of private-sector workers who were aged 35-45 at the time when a pension reform enacted in Italy changed the

[1] The old-age dependency ratio is defined as the ratio between the number of individuals aged 65 and above and the number of individuals in the working-age population (ages 20-64).

[2] As an example, the 1983 amendments to the US Social Security Act gradually raised the normal retirement age from 65 to 67 for cohorts born after 1937; affected individuals where therefore not older than 45 at the time of the announcement.

way in which a portion of their retirement benefits would be calculated. In 1995 the *Dini* pension reform, named after the then Prime Minister, started the transition to a notional defined-contribution (NDC) public pension scheme, whilst retaining a pay-as-you-go system. The reform affected the expected level of retirement benefits, likely lowering them for those who planned to claim early, as well as the ability to forecast their amount, by making them subject to macro-economic and demographic risks. Furthermore, it tightened the link between contributions paid to Social Security and benefits received; it also changed the incentive to delay claiming benefits past eligibility, thus possibly lengthening the working horizon.

The reform envisaged a long transitional phase and featured grandfathering clauses. Specifically, those who had accumulated at least 18 years of contributions to Social Security by 1995 were completely grandfathered: their retirement benefits would still be computed according to the pre-reform formula. On the other hand, those with less than 18 years of contributions were only partially grandfathered: the new NDC formula would apply to contributions accrued from 1996 onward, while the pre-reform formula would apply to earlier insurance years. I leverage this feature of the reform by comparing individuals who were barely fully grandfathered, *i.e.* those who had slightly more than 18 years of contributions (*control*), to those who were barely only partially grandfathered, *i.e.* those with slightly less than 18 years of contributions (*treated*).

I follow individuals from the beginning of the decade when the reform took place (1990) until 2006, when they are aged 46-56 and are still at least 4 or 5 years away from reaching benefit claiming eligibility. To do so, I exploit the availability of individual-level administrative panel data spanning the entire contribution histories of a random sample of private-sector workers. I leverage this data both to identify *treated* and *control* workers and to build measures of labour supply, by looking at days covered by (all and work-related only) contributions to Social Security, as well as the associated earnings. I estimate a generalized version of a difference-in-differences model. This specification allows me to asses both whether *treated* and *control* individuals were on parallel trends before the reform, as well as how post-reform labour supply adjustments unfold over time.

The setting provided by the *Dini* reform is attractive insofar as there are middle-aged individuals who learn years before coming near the retirement and benefit claiming decisions that there will be changes to the way in which a part of their pension benefits will be calculated. At the same time, there are similar individuals, those who just happen to have made slightly more contributions to Social Security by the end of 1995, who are not affected by new provisions. On the other hand, the setting has also some disadvantages. First, the reform was passed in August of 1995 and the variable that separates fully grandfathered to partially grandfathered individuals is not a fixed characteristic, such as the year of birth as in many reforms; it is instead a choice variable. Besides focusing on individuals with very similar contribution histories, in the main specification I exclude those who are in the 4-week window on either side of the 18-year threshold.

Second, as I explain in detail in Section 2.3, in 1992 there was another pension reform (the *Amato* reform) that on average affected to a larger extent *treated* individuals than *control* ones, by lengthening the reference period over which pensionable earnings were computed and by increasing the contribution requirement to claim an old-age pension. While pre-1995 coefficients in the generalized difference-in-differences specification show that *treated* and *control* individuals were on substantially parallel trends before 1995, I cannot claim that post-1995 coefficients only capture the impact of the *Dini* reform or that there would have not been any effect absent the *Dini* reform.

I find that *treated* individuals increase their labour supply by at most a small amount over the period studied; the effects become statistically significant few years after the *Dini* reform and tend to grow over time. For instance, the coefficients associated to the number of days covered by work-related contributions to Social Security, a measure of employment, is equal to 3 in 1998 and 5.5 in 2006, compared to an average of 292 among *treated* individuals in the pre-1996 period.

I then perform a set of robustness checks. First, I show that results remain very similar when including observations that are very close to the 18-year threshold. Second, I run placebo exercises whereby I set the threshold that separates fully grandfathered from partially grandfathered individuals at 19 or 20 years of contributions, rather than at 18. All individuals in the placebo samples are not affected by the changes in the computation of pension benefits brought about by the *Dini* reform. Post-1995 coefficients in the placebo exercises are in general small and not significantly different from 0. However, some notes of caution emerge as well.[3]

Finally, I perform heterogeneity analyses. When looking at each dimension of heterogeneity separately, I find that effects seem to be stronger among women and, by measuring education with a proxy, among high-educated individuals; furthermore, they tend to be larger among relatively low earners, who are more likely to see their retirement benefits reduced if they claim early. Results based on age at the time of the reform are inconclusive. The self-employed are more likely to see their benefits reduced than employees: while there is no substantial differential response in terms of days worked, self-reported earnings increase by more. I conclude by considering all dimensions of heterogeneity together: coefficients preserve their sign, but in most

[3] In the former exercise (19-year threshold), some of pre-1995 coefficients are significantly different from 0, suggesting a deviation from parallel trends before the *Dini* reform; in the latter exercise (20-year threshold) post-1995 coefficients relative to earnings are sometimes not much smaller than the true ones.

cases attenuate in size and lose statistical significance.

The remainder of the paper is organized as follows. Section 2.2 reviews the related literature; Section 2.3 explains the institutional setting; Section 2.4 describes the data; Section 2.5 presents the empirical strategy; Section 2.6 discusses the results; Section 2.7 concludes.

2.2 Related literature

This paper relates to the literature that studies how the design of public pension systems affects labour supply and that exploits pension reforms as natural experiments: due to grandfathering clauses and gradual phase-in schedules, many pension reforms allow the identification of groups of otherwise similar individuals who are subject to different retirement rules.

A seminal example is the work of Krueger and Pischke (1992). They examine the 1977 amendments to the US Social Security Act, which unexpectedly and substantially reduced social security benefits for the "notch" cohorts born in 1917-1921, while not affecting earlier cohorts. They find that lower social security wealth did not halt the declining trend in labor force participation of older individuals observed in the 1970s and 1980s. A number of papers have studied the effect of increasing the normal retirement age. The focus is mostly on employment, retirement and benefit claiming decisions of individuals "at risk of retirement" (Hanel and Riphahn (2012), p. 720), *i.e.* who are very close or past the eligibility for early retirement. The general finding is that a higher normal retirement age leads to delayed labour market exit and benefit claiming.[4] Other studies examine increases in the early retirement age. In these cases, the focus is mostly on employment and retirement effects at ages in between the old and the new early retirement age. These works generally report an increase in labour supply at these ages, but also provide evidence of program

[4]An example is Mastrobuoni (2009). He analyzes the 1983 amendments to the US Social Security Act, which gradually raised the normal retirement age from 65 to 67 for cohorts born after 1937: he focuses on individuals aged 61-65 - being 62 the early retirement age - and documents that the age at retirement increases by around half as a much as the increase in the normal retirement age. The same reform has been evaluated by Behaghel and Blau (2012); looking at individuals aged 62 and above, they report that the spike in benefit claiming moves from age 65 to the cohort-specific full retirement age. Another example is the work of Hanel and Riphahn (2012), who examine a reform that gradually raised the normal retirement age for women from 62 to 64 in Switzerland, while introducing a penalty for early benefit claiming at age 62. Focusing on women aged 60-65, they find a decline in retirement probability at the pre-reform normal retirement age. The same reform has been analyzed in the work of Lalive et al. (2017).

substitution effects.[5,6]

Closer to my work, a number of papers study whether pension reforms affect also the behaviour of individuals who are farther away from retirement and benefit claiming decisions. French (2005) develops a model of life-cycle labour supply, retirement and saving behaviour where individuals can save to buffer against health shocks and wage shocks, as well as to finance retirement, whereas they cannot borrow against future income, including that from Social Security; furthermore there are fixed costs of working on the sides of the employees and the employer. In such a model, most life-cycle labour supply adjustments would occur along the extensive margin; this implies that the labour supply elasticity of older workers, who are nearest the participation margin decision, is higher than that of younger cohorts. Hence, retirement incentives would affect the former to a larger extent than the latter. For instance, simulations show that a 20 per cent reduction of social security benefits, framed within the US Social Security and pension system, would increase the labour supply at all ages, but especially past the early retirement age (age 62; French, 2005 and French and Jones, 2012).

Duggan et al. (2007) examine the effect of the 1983 amendments to the US Social Security Act on enrollment into the disability insurance program: they find an increase in take-up rates not only among individuals aged 55-64, but also among younger cohorts (of ages between 45 and 54). Other papers evaluate the effect of changes to the minimum retirement age, which lengthen the working horizon of individuals. Exploiting changes that took place in Italy between the mid-1990s and the early 2000s, Brunello and Comi (2015) report that a higher minimum retirement age increases participation in training programs among private-sector workers aged 40-54.[7] Leveraging a pension reform that in 2004 increased the age requirement to claim a seniority pension from 57 (resp. 58) to 60 (resp. 61) for employees (resp. self-employed), Bertoni et al. (2018) further show that an expected longer working horizon increases healthy behaviours among Italian male workers of ages between

[5]Examples are the works of Atalay and Barrett (2015) and Cribb et al. (2016). The former studies the effect of gradually extending the female minimum age for claiming the age pension in Australia from 60 to 65. Restricting the attention to women aged 60-64, it reports a 12 to 19 per cent decline of the probability of retirement per every one-year increase of age eligibility. The latter assesses the effects of gradually increasing the early retirement age for women in UK from 60 to 62, documenting a positive impact on employment at ages 60 and 61 (6.3 percentage points). Another example is the work of Staubli and Zweimüller (2013), who evaluate the increase of the early retirement age in Austria, from 60 to 62 for men and from 55 to 58.25 for women. They study employment effects in the broader age range 57-64 for the former and 52-59 for the latter: they document that a one year increase of the early retirement age raises employment of women and men during that year by 11 and 9.75 percentage points, respectively.

[6]Other recent papers that study the effect of public pension systems on labour supply are Liebman et al. (2009), Vestad (2013), Fetter and Lockwood (2018) and Seibold (2019).

[7]Montizaan et al. (2010) find similar effects when studying participation on training courses at age 56 for Ducth public-sector employees who were subject to the abolition of pre-pension plans that would have allowed them to retire at age 62.25; the same workers also experienced a deterioration of mental health at age 58 (de Grip et al., 2011).

42 and 51.

Focusing on labour supply, Hairault et al. (2010) exploit the features of the French pension system and a reform that in 1993 increased the number of insurance quarters needed to claim a pension. They study how the distance to retirement influences the employment probabilities of workers aged 50-59, 60 being the minimum retirement age: they find a positive effect among workers aged 55-59, but not among younger cohorts. Leveraging a pension reform that suddenly and heterogeneously changed the minimum retirement age in Italy in 2011, Carta and De Philippis (2019) document an increase in labor force participation not only for women aged 55-59, but also - albeit to a lower extent - for younger ones (of ages between 45 and 54); they call this effect a "perspective effect".[8] The "anticipation effect" is also studied in Engels et al. (2017), who analyze the 1992 German pension reform that introduced deductions when claiming benefits between the early retirement age (60) and the normal retirement age. They report that employment rose not only for women aged 60 and above, but also for women aged 55-59.[9]

Finally, there are papers which have studied the effect of the Italian pension reforms that took place in the 1990s on the accumulation of private wealth. Attanasio and Brugiavini (2003) examine the 1992 *Amato* pension reform, which reduced social security wealth to a larger extent for public sector employees than for private sector ones and for younger cohorts than for older ones: based on data that spans the 1989-1995 period, they document an increase in private saving rates. As in my work, Bottazzi et al. (2006) consider differences among workers with more or less than 18 years of contributions by the end of 1995, the threshold that in the 1995 *Dini* reform separates those fully grandfathered from those only partially grandfathered. They find that individuals partially update their expectations of age at retirement and the replacement rate following the reform of the 1990s. Furthermore they confirm the positive effect on the accumulation of private wealth, particularly among workers who are better informed about changes brought about by the reforms; the component that increases more is real estate wealth (Bottazzi et al., 2011).

2.3 Institutional setting

In 1992 public pension spending in Italy, which runs a pay-as-you-go (PAYG) system, amounted to 12.5 per cent of GDP, almost two times as large as the OECD average. This figure reflected a combination of factors: a high old-age dependency ratio (25); generous rules governing both the computation of pension benefits and eligibility for early and full retirement, especially for public sector employees.

[8]Before the reform, private sector female employees became eligible for an old-age pension upon turning 60.

[9]They also find that unemployment decreases before age 60, while increasing after that age.

In September 1992 Italy experienced a foreign-exchange crisis; in 1993 real GDP fell by 0.7 per cent, recording the first contraction since 1975, and employment decreased by 2.8 per cent, the largest drop since the post-war period (Banca d'Italia, 1993). Against this backdrop of deteriorating macroeconomic fundamentals and high pension spending, two reforms of the pension system were implemented: the 1992 *Amato* reform and the 1995 *Dini* reform. Both reforms envisaged grandfathering clauses and long transitional periods before reaching a complete phase-in. Due to this, there are otherwise similar workers who have been affected by new provisions to a different extent.

The existence of full and partial grandfathering clauses created a web of complex rules for the computation of benefits during the transitional phases. Nonetheless, it is always possible to re-write the formula that computes first year annual pension benefits (b) as a generalization of the following expression:

$$b = \text{yield rate} \cdot \text{pensionable earnings} = \gamma \cdot f(y_{t_h}, y_{t_{h+1}}, ..., y_{R-1}, y_R; \Pi) \qquad (2.1)$$

where, borrowing the terminology adopted by Hamann (1997), γ and Π are the yield rate and the accrual rate, respectively. Earnings y from year t_h to the last contribution year R are re-evaluated according to the accrual rate Π and transformed into pensionable earnings based on a certain function f. Pensionable earnings are then converted into pension benefits using the yield rate γ.[10] The *Amato* reform lengthened the reference period $(t_h : R)$ over which to compute pensionable earnings. Within the framework of a PAYG system, the *Dini* reform started the transition to a notional defined-contribution (NDC, henceforth) regime, which also implied changes to γ, f and Π.

The following subsections describe the main features of the pension system before the reforms and review the main regulatory changes for private sector workers, who are the focus of the analysis in this paper. I distinguish, where applicable, provisions which affected all individuals in a similar way from those that had a heterogeneous impact due to grandfathering clauses. Further details on the formulae that are used to compute pension benefits can be found in Appendix Section 2.A.

2.3.1 The pre-1993 regime

Female (resp. male) employees became eligible for an old-age pension upon turning 55 (resp. 60) years old, conditional on having 15 years of contributions to Social Security (SS contributions, henceforth).[11] 35 years of SS contributions, regardless of age, entitled workers to a seniority pension. First year annual pension

[10] The accrual rate and the yield rate are therefore the rates used to re-evaluate past earnings and to convert pensionable earnings into first year pension benefits, respectively.

[11] For female and male self-employed the age requirement for an old-age pension was being at least 60 and 65 years old, respectively.

benefits were computed according to the following version of equation (2.1):

$$\begin{aligned} b &= k \cdot min(N_{t_0:R}, 40) \cdot \bar{y}_{t_l:R} \\ &= k \cdot min(N_{t_0:R}, 40) \cdot \frac{1}{(L/52)} \cdot \left(\sum_{t=t_l}^{R} y_t \Psi_t \right) \end{aligned} \tag{2.2}$$

with:

$$\Psi_t = \prod_{j=t}^{R} (1+\pi_j) \text{ for } t = t_l, ..., R-2 \text{ and } \Psi_{R-1} = \Psi_R = 1 \tag{2.3}$$

$$t_l : \sum_{j=t_l+1}^{R} \frac{n_j}{52} < \frac{L}{52} \leq \sum_{j=t_l}^{R} \frac{n_j}{52} \tag{2.4}$$

where n_j are weeks covered by SS contributions in year j. Pensionable earnings were computed according to a complex function f. Earnings made in the final L weeks covered by SS contributions were: re-evaluated according to the accrual rate Ψ (*coefficiente di rivalutazione*), which reflected inflation (π); averaged ($\frac{1}{L/52}$); and multiplied by lifetime years of SS contributions, which were however capped at 40 ($min(N_{t_0:R}, 40)$). L was equal to 260 weeks (5 years) for employees and to 520 weeks (10 years) for the self-employed.[12]

k, the yield rate (*aliquota di rendimento*), was 2 per cent for individuals whose average earnings in the final L weeks covered by SS contributions were below a certain threshold. The threshold was adjusted upward every year to take into account inflation. The average yield rate then declined, due to the fact that the marginal yield rate was a decreasing step-way function of average earnings. This implied a lower replacement rate for high-earners. Pension benefits for low-earners, on the other hand, could amount to as much as 80 per cent of their final L-week average earnings (2 per cent $\times$ 40 years of SS contributions).

While in the formula outlined in (2.2) pension benefits are a function of earnings and not of money paid as SS contributions, years covered by SS contributions concur to determine pensionable earnings, thus making the entire contribution history up to the fortieth year relevant for the computation of benefits. The fact that years of SS contributions beyond the fortieth were not taken into account lowered the incentive to continue working past that date.

Pension benefits after the first year were indexed by taking into account both the inflation rate and the real growth rate of earnings.

[12]Notice that if there are gaps in the contribution history of an individual, the final 260 (520) weeks covered by SS contributions do not coincide with the last 260 (520) calendar weeks. In year t_l only the portion of earnings that is made during the weeks that are needed to reach the 260 (520) threshold is included in the summation term of (2.2).

2.3.2 The 1992 *Amato* reform

The 1992 *Amato* reform intervened both on eligibility criteria for old-age pensions and on the computation of pension benefits.

2.3.2.1 Provisions that affected all individuals in a similar way

The age requisite for old-age pensions of private sector employees was gradually increased, over the period 1993-2000, from 55 to 60 for women and from 60 to 65 for men. Pension benefits after the first year were indexed only to inflation. First year annual pension benefits became the sum of two "quotas", Quota A and Quota B:

$$b = \underbrace{k_A \cdot N_A \cdot \bar{y}_{t_l:R}}_{\text{Quota A}} + \underbrace{k_B \cdot N_B \cdot \bar{y}_{t_q:R}}_{\text{Quota B}} \tag{2.5}$$

The yield rate of Quota A (k_A) was the same as before; with respect to pensionable earnings, the only difference was that the average final L-week annual earnings were no longer multiplied by lifetime years of SS contributions, but only by years of SS contributions up to 1992 (N_A).

With regards to Quota B, the yield rate (k_B) started at 2 per cent as for Quota A, but its decline at higher levels of average earnings followed a slightly different trajectory than that of k_A.[13] The years of SS contributions that concurred to determine pensionable earnings were those from 1993 to retirement (N_B). If total years of SS contributions ($N_A + N_B$) exceeded 40, the excess years beyond the fortieth were subtracted from the computation of the "quota" that featured the lowest average earnings. Furthermore, earnings were re-evaluated according to a higher accrual rate Ψ_B, that took into account inflation and added a fixed 1 per cent for every year until retirement.[14]

2.3.2.2 Provisions that affected workers differently due to grandfathering clauses

With regards to average earnings in the Quota B part of pension benefits ($\bar{y}_{t_q:R}$), the *Amato* reform aimed at extending the reference period to the entire contribution life. On this matter, however, the reform envisaged different grandfathering clauses for three group of workers: (i) those with at least 15 years of SS contributions by 1992 (T1-92; mostly grandfathered); (ii) those with less than 15 years of SS contri-

[13]For those who started collecting pension benefits in 1993, for instance, the marginal yield rates of Quota A were the following: 2 per cent in the [0, 53,475,000] bracket; 1.50 per cent in the (53,475,000-71,121,750] bracket; 1.25 per cent in the (71,121,750; 88,768,500] bracket; 1 per cent above 88,768,500. The marginal yield rates of Quota B were the following: 2 per cent in the [0, 53,475,000] bracket; 1.60 per cent in the (53,475,000-71,121,750] bracket; 1.35 per cent in the (71,121,750; 88,768,500]; 1.10 per cent in the (88,768,500-101,602,500] bracket; 0.9 per cent above 101,602,500. Money amounts are expressed in Italian Liras.

[14]Appendix Section 2.A provides the formal expressions of N_A, N_B and Ψ_B.

butions by 1992 (T2-92; partially grandfathered); (iii) those who started accruing contributions from 1993 onward (T3-92; not grandfathered).

Specifically, for individuals in the T1-92 group the reference period was gradually lengthened to encompass the last 520 (resp. 780) weeks covered by SS contributions for employees (resp. self-employed). For workers belonging to the T2-92 group the reference period was further extended, to include all weeks covered by SS contributions from 1993 to retirement and the last 260 (resp. 520) weeks covered by SS contributions before 1993 for employees (resp. self-employed). Finally, for individuals who entered the labour market after 1992 (T3-92) there was no Quota A, and the reference period of Quota B spanned the entire contribution history.

Furthermore, the contribution requisite for claiming an old-age pension was gradually raised from 15 years to 20 years over the 1993-2001 period for most workers, but those belonging to group T1-92 were grandfathered.[15]

2.3.3 The 1995 *Dini* reform

The 1995 *Dini* reform intervened both on eligibility criteria for early retirement and, in a more radical way than the *Amato* reform, on the method to compute first year pension benefits. All these changes affected workers differently depending on years of contributions by 1995.

Specifically, the reform legislated that SS contributions would accrue into a notional account where they would be capitalized according to the 5-year moving average of nominal GDP growth rate (g) and then summed to form the individual's notional capital. The accumulated capital would then be converted into first year annual pension benefits by applying a transformation coefficient (β_a, *coefficiente di trasformazione*), which is an increasing function of the benefit claiming age, as it reflects the lower residual life expectancy of individuals who start claiming benefits later. It was also legislated that the transformation coefficient had to be periodically adjusted to reflect the evolution of life expectancy. Assuming a constant social security tax rate (τ), the formula for computing first year annual pension benefits once the NDC regime is completely phased-in can be written as:

$$b = \tau \cdot \beta_a \cdot \left(\sum_{t=t_0}^{R} y_t G_t \right) \tag{2.6}$$

[15]The two other grandfathered groups were the following: first, those who had started accumulating contributions at least 25 years before 1992 and had worked less than a full year for at least ten years; second, those who had started accumulating contributions before 1992 and could not reach 20 years of SS contributions by the time they would have satisfied the age requisite for an old-age pension.

with:

$$G_t = \prod_{j=t}^{R}(1+g_j) \text{ for } t = t_0, ..., R-2 \text{ and } G_{R-1} = G_R = 1 \tag{2.7}$$

The social security tax rate is equal to 33 per cent for employees, while it has gradually risen from 20 to 24 per cent for the self-employed.[16] G is the accrual rate (*coefficiente di capitalizzazione*), which is a function of nominal GDP growth; $\tau \cdot \beta_a$ can be thought as the yield rate. The function f is simpler than in the Quota A and Quota B formulae, as pensionable earnings are just the sum of re-evaluated earnings from year t_0 to the last contribution year R.

The transition to the NDC regime was, however, very gradual. Specifically, the reform identified three groups of workers, who were subject to different grandfathering clauses: (i) those with at least 18 years of SS contributions by 1995 (C95); (ii) those with less than 18 years of SS contributions by 1995 (T1-95); (iii) those who started contributing after 1995 (T2-95). Individuals belonging to group C95 were fully grandfathered: pension benefits would still be computed using the formula modified by the Amato reform. Workers belonging to the T1-95 group were only partially grandfathered: the NDC system would apply *pro-rata* only to contributions accrued from 1996 onward. Finally, individuals entering the labour market after 1995 were fully covered by the new provisions.

In 2011 a major pension reform, the *Fornero* pension reform, extended the NDC method *pro-rata* (i.e. to SS contributions made after 2011) also to individuals belonging to group C95. However, since my analysis spans over a period that does not include that reform (1990-2006), I do not consider these further changes to the formulae used to compute pension benefits.

Therefore, abstracting from the *Fornero* reform, following the 1995 *Dini* pension reform social security benefits of individual i are computed in the following way:

$$b = \begin{cases} k_A \cdot N_A \cdot \bar{y}_{t_l:R} + k_B \cdot N_B \cdot \bar{y}_{t_v:R}, \text{if } i \in \text{C95} \\ k_A \cdot N_A \cdot \bar{y}_{t_l:R} + \\ +k_B \cdot N_{1993:1995} \cdot (\bar{y}_{t_v:R} \cdot \mathbb{1}(i \in \text{T1-92}) + \bar{y}_{t_w:R} \cdot \mathbb{1}(i \in \text{T2-92}) + \bar{y}_{t_0:R} \cdot \mathbb{1}(i \in \text{T3-92})) + \\ +\tau \cdot \beta_a \cdot \left(\sum_{t=1996}^{R} y_t G_t\right), \text{if } i \in \text{T1-95} \\ \tau \cdot \beta_a \cdot \left(\sum_{t=t_0}^{R} y_t G_t\right), \ \text{if } i \in \text{T2-95} \end{cases} \tag{2.9}$$

[16] For notational simplicity and to make the comparison with other formulae easier, τ is assumed to be constant in equation (2.6). With a time-varying social security tax rate, the NDC formula reads:

$$b = \beta_a \cdot \left(\sum_{t=t_0}^{R} \tau_t y_t G_t\right) \tag{2.8}$$

where the explicit expressions for $\bar{y}_{t_v:R}$, $\bar{y}_{t_w:R}$ and $\bar{y}_{t_0:R}$ are provided in Appendix Section 2.A. As explained before, for individuals with at least 18 years of contributions nothing changed with respect to what had been envisioned by the *Amato* reform. On the other hand, for individuals with less than 18 years of contributions, the weight of Quota B was greatly reduced, as it was only applied to contributions made between 1993 and 1995 ($N_{1993:1995}$).

For individuals entering the labour market after 1995, the seniority pension was abolished and retirement benefits could be claimed from age 57, conditional on having 5 years of *effective* contributions (see Section 2.4). For individuals belonging to groups T1-95 and T2-95 a seniority pension could be claimed upon reaching 57 years of age and 35 years of SS contributions, or upon reaching 40 years of SS contributions, regardless of age.[17]

2.3.4 How grandfathering clauses made similar workers different

After reviewing the main changes brought about by the 1992 and 1995 pension reforms, I conclude this section by describing how the existence of grandfathering clauses made otherwise similar workers subject to different regimes for the computation of pension benefits. I also discuss how such differences could affect labour supply decisions. I restrict the attention to individuals barely assigned to the C95 group (*i.e.* workers with 18 or slightly more than 18 years of contributions by 1995) and individuals barely assigned to the T1-95 group (*i.e.* workers with slightly less than 18 years of contributions by 1995). These workers will be the focus of the analysis carried out in the next sections.

The first thing to notice is that individuals belonging to the C95 group in 1995 were in the T1-92 group three years before.[18] Among individuals belonging to the T1-95 group, on the other hand, there are all individuals who are part of the T2-92 group, because workers with less than 15 years of SS contributions by 1992 could not have at least 18 years of SS contributions by 1995, and some workers - those who had large enough contribution gaps between 1993 and 1995 - who belonged to the T1-92 group. Figure 2.1 depicts this graphically.

The 1992 *Amato* reform therefore affected individuals belonging to groups C95 and T1-95 in a different way: first, the reference period over which to compute pensionable earnings was lengthened on average to a lower extent for the former than for the latter. The degree to which an extended reference period affects the level of first year pension benefits depends on the steepness of the earnings-age profile: given that

[17]The age requisite for the seniority pension with 35 years of contributions and the contribution requisite for the seniority pension with 40 years of contributions were phased-in gradually, over the period 1996-2008.

[18]Workers who had at least 18 years of contributions by 1995 also had at least 15 years of contributions by 1992.

earnings in the Quota B formula are re-evaluated based on the inflation rate and a fixed 1 per cent premium, including earlier earnings would increase pensionable earnings for individuals who experience a null or feeble real earnings growth, whereas it would reduce pensionable earnings for individuals with a more dynamic career. While the incentive to work and contribute to social security throughout an individual's life was already present for everybody in the pre-1992 formula (due to the presence of the term $min(N_{t0:R}, 40)$), extending the reference period for computing pensionable earnings could increase the incentive to report high earnings earlier in life. This could be especially relevant for the self-employed, who, unlike employees, are not subject to double-reporting of earnings by third parties, and might therefore more easily under-report them.

Second, the contribution requirement to claim an old-age pension increased more for individuals in the T1-95 group (from 15 to 20 insurance years). A higher contribution requirement can result in increased labour supply, in order to meet it. However, since I will focus my analysis on individuals who already have slightly more or less than 18 years of SS contributions by 1995, this further difference may not have a major differential effect on labour supply decisions.

The 1995 *Dini* reform then generated further and larger differences among the two groups of individuals: while it did not affect individuals belonging to the C95 group, it further changed the way in which pension benefits are computed for individuals in the T1-95 group. Hamann (1997) provides a discussion of the long-run properties of the *Amato* and the *Dini* systems, in a scenario where they would have both completely been phased-in. I elaborate on it and adapt it to assess the implications of having SS contributions made from 1996 onward contribute to determine first year pension benefits according to the NDC formula, rather than through the Quota B formula.

First, for individuals belonging to the T1-95 group the level of pension benefits can change. As explained in Subsection 2.3.3, the *Dini* reform has changed the yield rate ($\tau\beta_a$ as opposed to k_B) and the accrual rate (G as opposed to Ψ_B). Furthermore, it has changed the function f, i.e. the way in which earnings over the period $(t_h : R)$ are transformed into pensionable earnings, as well as the starting year of the reference period (t_h).

Figure 2.2 compares the accrual rate in the Quota B and in the NDC regimes. On the x-axis, there is the year when one euro is paid into SS contributions (contribution year), while on the y-axis there is the value of this euro in the year when the individual starts claiming benefits (first benefit claiming year). As I will focus on individuals aged 35-45 at the time of the *Dini* reform, and given the eligibility criteria for old-age and seniority pensions prevailing at that time, the first benefit claiming years span from 2011 to 2026.[19] Figure 2.2 shows that the accrual rate is higher in the NDC regime for earlier contribution years and in the earliest first benefit claiming years; this is due to the relatively good performance of nominal GDP growth in the 1990s, apart from the 1992-1993 crisis. For individuals retiring later, the fewer are the contribution years for which the accrual rate is higher in the NDC system and the larger is the difference in favour of the Quota B accrual rate. This shows that periods of subdued macro-economic performance, such as the one that started following the 2007-2008 financial crisis and the 2011 sovereign debt crisis, negatively affect the level of pension benefits in the NDC system. In my analysis, I focus on labour supply responses over the 1996-2006 period: the 5-year moving average of the real GDP growth rate peaked at the beginning of the 2000s and then started to decline. The extent to which individuals could foresee that the performance of the Italian economy would further worsen is unknown and likely varies based on individual characteristics such as education and occupation.

Figure 2.3 compares the yield rate in the Quota B and in the NDC regimes.[20] The figure features two panels because employees and self-employed have different yield rates in the NDC regime.[21] It shows that for employees, the yield rate under the NDC regime is higher than the one under the Quota B regime only for high-earner individuals (*i.e.* those for whom the marginal yield rate in the Quota B regime is below 2 per cent) who start collecting retirement benefit relatively late (as the transformation coefficient in the NDC regime increases with the benefit claiming age). For the self-employed, given the reduced social security tax rate, the yield rate of the NDC regime is always lower.

Overall, given the performance of the Italian economy, individuals belonging to the T1-95 group are likely to experience a decline in first year annual pension benefits if they are not high-earners and if they start collecting retirement benefits early; the drop could be particularly pronounced for the self-employed, due to the lower yield rate. On the other hand, high-earners and individuals who retire late could benefit from the transition to the NDC regime, in which the yield rate increases with age at

[19]This interval is obtained assuming that individuals would start claiming benefits on January 1st of the year after the first one in which they reach eligibility for either an old-age or a seniority pension, according to rules prevailing at the time of the *Dini* reform.

[20]The yield rate in the NDC regime is an increasing function of the benefit claiming age and was revised downward in 2010, 2013, 2016 and 2019 to reflect increases in life expectancy. The yield rate in the Quota B regime does not vary with the benefit claiming age, but it starts at 2 per cent and then declines as average earnings grow and surpass various thresholds. In the figure I plot the yield rate for average earnings below the first threshold, as well as the average yield rate for average earnings that equal exactly the following thresholds. As an example, let the marginal yield rate be 2 per cent in the bracket $[0, x_1]$ and 1.6 per cent in the bracket $(x_1, x_2]$. The average yield rate (k_B) for an individual whose average earnings ($\bar{y}_{t_q:R}$) are equal to x_2 is found by solving the equation $k_B \cdot N_B \cdot x_2 = 0.02 \cdot N_B \cdot x_1 + 0.016 \cdot N_B \cdot (x_2 - x_1)$.

[21]The social security tax rate for the self-employed has gradually increased from 20 to 24 per cent; the yield rate in Figure 2.3 is computed assuming a social security tax rate equals to 24 per cent.

retirement. Furthermore, there are two features of the Quota B formula that penalize them: the yield rate which declines as average earnings grow and the 40-year cap on years of SS contributions that can be taken into account. In Appendix Section 2.B, I describe how I simulate the expected change in annual pension benefits experienced by individuals barely assigned to the T1-95 group. I find that across several alternative scenarios, whereby I vary the assumptions about real earnings growth, the number of days worked and retirement decisions, the majority of individuals in my sample (see Section 2.6) would experience a negative change.

Second, the transition to the NDC regime affects an individual's ability to forecast the level of pension benefits. In the Quota B regime the accrual rate is a function of inflation and a fixed 1 per cent premium. In the NDC regime, on the other hand, it is a function of the nominal GDP growth rate: besides reflecting changes in the cost of living, it depends on the performance of the Italian economy and is therefore subject to macro-economic risks. With regards to the accrual rate, while it is fixed (within a given income bracket) in the Quota B regime, it is periodically adjusted to reflect changes in life expectancy in the NDC regime and is thus subject to demographic risks.

Third, in the NDC regime the link between SS contributions and pension benefits is tighter. In the Quota B regime only years covered by SS contributions, but not the amount paid, matter for the computation of pension benefits. In the NDC regime, on the other hand, pension benefits are a function of money paid as SS contributions.

Fourth, the 1995 reform affects the incentives to delay collecting retirement benefits past eligibility. In the Quota B regime, an additional year of work increases first year pension benefits insofar as it translates into higher pensionable earnings. However, the yield rate does not vary with the benefit claiming age; furthermore, years of contributions beyond the fortieth year are not taken into account, thus creating an implicit disincentive to work past that date. In the NDC regime, on the other hand, the transformation coefficient is a positive function of the benefit claiming age, which increases the incentives to work an additional year and delay labour market exit.

The changes described above can affect through multiple channels the labour supply decisions of individuals in their prime-age and, in general, of those not yet near to retirement eligibility. The *Amato* and *Dini* pension reforms likely generated income effects, insofar as they affected the level of pension benefits at any given benefit claiming age. If most individuals belonging to the T1-95 group expect lower pension benefits, which would typically be the case for the self-employed and for individuals in the low and middle part of the earnings distribution planning to retire relatively early, in order to maintain adequate levels of income when old they may increase labour supply not only near or past retirement eligibility, but also during middle-age. The fact that in the NDC regime the transformation coefficient is an increasing function of age at retirement and that years of contributions after the fortieth year are not discarded also generates a price effect, increasing the incentives to continue working an additional year. A longer working horizon may induce individuals to increase labour supply also during prime-age. A number of reasons are discussed by Carta and De Philippis (2019). They find that an extended working horizon, due to an increase of the full retirement age brought about by the 2011 *Fornero* pension reform, has a positive effect on the labour supply of middle-aged Italian women: by increasing the expected duration of a job, a longer working horizon raises the value of searching for it while younger and then retaining it; individuals may prefer to supply more labour at younger ages rather than at older ages, if the probability of experiencing negative health or productivity shocks and the cost of working increase with age.

Furthermore, the increase in uncertainty about the expected level of pension benefits could induce individuals to increase savings for "precautionary" motives.[22] One way to save more is to increase labour supply. Giavazzi and McMahon (2012), for example, find that German household heads who worked part-time extended working hours in the period around the close 1998 election, a time that was characterized by uncertainty, among other things, as to whether a pension reform that had lowered pension benefits would have been confirmed or revoked. Finally, the tighter link between contributions paid and benefits received could reduce the distortionary effect of Social Security taxation on labour supply decisions, if individuals perceive contributions not as pure taxes but as deferred earnings (*e.g.* Auerbach and Kotlikoff (1985), Hamann (1997)).

2.4 Data

I leverage and link on account of unique individual identifiers administrative data provided by the Italian Institute of Social Security (*Istituto Nazionale della Previdenza Sociale*, INPS). The main dataset contains full contribution histories for a random sample of Italian private-sector workers. Auxiliary demographic information (gender, year and country of birth) is recorded in the workers register. The unit of observation in the main dataset is the triple individual-contribution spell-year. The information provided is the following: (i) the start and the end dates of the spell; (ii) the length of the period, which can be expressed in days, weeks or months, covered by SS contributions; (iii) the event (*e.g.* paid work, short-time work, maternity leave) and the earnings associated to the contribution spell; (iv) the pension fund contributions are paid into (*e.g.* the fund for private sector employees, the fund for

[22] Carroll and Kimball (2008) provide a review of the literature on precautionary savings.

self-employed).

The events that lead to the accrual of SS contributions are of two types: *effective* and *figurative*. *Effective* events consist of paid work, on which the Social Security tax is levied. *Figurative* events are circumstances of suspended or reduced work activity, which nonetheless result in contributions being credited to workers' accounts; the main *figurative* events are maternity, sickness or work-related injury, unemployment and short-time work.

The importance of observing full contribution histories is twofold. First, it allows the identification of workers belonging to the groups T1-95 and C95, by computing years of SS contributions accumulated by the end of 1995 (see Appendix Section 2.C for details). Second, it allows the construction of various measures of labour supply (see Appendix Section 2.D for details). Specifically, I build five measures. The first two are proxies for labour market participation: (i) a dummy that takes value 1 if at least one day in the year is covered by SS contributions; and (ii) the number of days covered by SS contributions per year.[23] The other three are meant to measure employment: (iii) a dummy that takes value 1 if at least one day in the year is covered by work-related SS contributions, as can be inferred from the event associated to the contribution spell; (iv) the number of days covered by work-related SS contributions per year; and (v) yearly labour earnings, expressed in 2016 euros.[24]

2.5 Empirical strategy

On the one hand, the institutional context described in Section 2.3 provides an attractive setting to study whether changes to the public pension system affect the labour supply of individuals who were in their prime-age when they learned about them: following the 1995 *Dini* pension reform, otherwise similar workers - those with slightly less or more than 18 years of SS contributions by the end of 1995 - are subject to different regimes for the computation of pension benefits.

On the other hand, some features of the setting make the identification of causal effects challenging. First of all, the extent to which workers are affected by the *Dini* reform does not depend on a fixed characteristic, such as the date of birth, but on a choice variable - years of SS contributions by 1995. This also implies

[23]Contributions related to the redemption of periods spent in the army or in higher education are not considered, although estimates are virtually unchanged when they are taken into account. These variables are just a proxy of labour market participation because, according to the ILO definition, an individual is defined as active in the labour market if he/she has worked or searched for a job in a reference period. The information contained in contribution histories does not allow to build this exact measure of labour market participation; as an example, an individual who is searching for a job but is not receiving any kind of social insurance benefits would be classified as inactive.

[24]For some individual-contribution-spell-years the information on labour earnings is missing. This explains the lower number of observations in regressions presented in Section 2.6.

that the assignment to the T1-95 (*treated*, henceforth) or C95 (*control*, henceforth) groups depends on the lagged realizations of the outcome that I wish to study - labour supply. This complicates the task of separating the effect of the reform from unrelated differences in post-reform labour supply trajectories of individuals who supplied a different amount of labour in the pre-reform period.

I attempt to mitigate this concern by focusing on individuals aged 35-45 in 1995 with a similar contribution history up to that year: I compare individuals barely assigned to the *control* group, which I define as those with 18 (included) to 19 years of SS contributions by the end of 1995, to individuals barely assigned to the *treated* group, which I define as those with 17 to 18 (excluded) years of SS contributions by the end of 1995.[25] Table 2.1 reports some relevant statistics, as of 1995. Individuals in the *treated* and *control* groups are very similar in terms of demographics. The former are only 2 percentage points less likely to be male and are slightly more than one month younger. They are as likely as the latter to be born in Italy and to work as self-employed. On the other hand, individuals in the *treated* group contributed less and worked less in 1995, as measured by days covered by SS contributions and by days covered by work-related SS contributions, respectively. As a consequence, their earnings were lower. These differences have to be taken into account.

To this end, I take advantage of the availability of individual-level panel data to estimate the following generalized version of a difference-in-differences model with individual fixed effects:

$$y_{it} = \lambda_i + \gamma_t + \sum_{k=t_0}^{t_1} \beta_k T_i \cdot \mathbb{1}(\text{year} = k) + \varepsilon_{it} \tag{2.10}$$

i indexes the worker and t indexes the year; y_{it} is one among the various measures of labour supply (see Section 2.4); T_i is a dummy variable taking value 1 if individual i belongs to the *treated* group. λ_i and γ_t are individual and year fixed effects, respectively. Individual fixed effects allow control for time-invariant sources of heterogeneity across workers; year fixed effects further account for year-specific shocks that affect all individuals. ε_{it} is the error term. Standard errors are clustered at the individual level. β_k are the coefficients of interest; I set β_{1995} equal to 0, so that a given coefficient β_k measures how different the labour supply is of individuals belonging to the *treated* and *control* groups in year k relative to year 1995, when the difference is normalized to 0.

[25] I exclude individuals who have contribution spells before 1996 in pension funds that are not those of private sector employees and self-employed, as different rules apply. This leads to discarding 10.6 per cent of observations. As my sample period spans from 1990 to 2006, I also drop individuals who die before 2007, which amounts to 2.3 per cent of the remaining observations. Finally, I discard workers who receive an old-age or seniority pension before 2007: this leads to a drop of only 0.27 per cent of observations, as individuals aged 35-45 in 1995 and with 17 to 19 years of contributions were in principle not eligible to obtain those types of pensions by 2006.

I follow workers from the beginning of the decade when the pension reforms take place (1990, t_0) to 2006 (t_1). During this period, individuals aged 35-45 and with 17 to 19 years of contributions by 1995 were not eligible to collect retirement benefits: in 2006 women (resp. men) could claim an old-age pension upon turning 60 (resp. 65) years old; a seniority pension could be obtained upon accumulating at least 35 years of contributions, and conditional on being at least 57 years old, or upon accumulating 39 years of contributions regardless of age.[26] I, therefore, study labour supply responses during a period that substantially coincides with the middle-age phase of life and terminates when individuals are still at least 4 or 5 years away from gaining eligibility to collect pension benefits.

A second issue is that, as explained in Subsection 2.3.4, individuals belonging to the *treated* and the *control* groups were already affected to a different extent by the 1992 *Amato* reform. The visual inspection of coefficients $\{\beta_k\}_{k=1990}^{k=1994}$ in Section 2.6 will show that for most outcomes they are not significantly different from 0, meaning that individuals in the *treated* and *control* groups appear to have been on substantially parallel trends in terms of labour supply in the years preceding the *Dini* reform. This, however, does not imply that the *Amato* reform would not have had an effect, had the *Dini* reform not been passed, or that differences in labour supply decisions after 1995 are only due to the provisions of the *Dini* reform. This has to be taken in mind when interpreting the results. I also estimate the following more compact difference-in-differences specification:

$$y_{it} = \lambda_i + \gamma_t + \beta T_i \cdot Post_t + \varepsilon_{it} \tag{2.11}$$

where $Post_t$ is a dummy variable that takes value 1 in the years following the *Dini* reform.

The *Dini* pension reform was passed in August 1995. This raises the concern that individuals could already change their behaviour in the fall of 1995 in order to end up on the preferred side of the 18-year threshold. Figure 2.4 plots the density histogram of years of SS contributions by the end of 1995 for the sample of individuals who were aged 35-45 at that time and who had accumulated 14 to 22 years of contributions. The spike observed at around 18 years of contributions is not of a substantially different magnitude with respect to the spikes observed around other integer years of contributions. These spikes are likely due to the fact that many job relationships start in January, so that a disproportionate share of individuals will have an integer number of years of contributions by the end of any given year. Nonetheless, I attempt to mitigate this concern by estimating as the preferred specification a "donut" difference-in-differences specification (Baltrunaite et al., 2018),

[26] The age requirement for the seniority pension with 35 years of contributions was 58 for self-employed.

whereby I exclude workers lying in the 4-week window on either side of the 18-year threshold.[27] In a robustness check I will show that results are however very similar when including those individuals.

After presenting the main results on the entire sample, I look whether labour supply responses are heterogeneous depending on individual socio-demographics characteristics. To this end, let $d = \{0, 1\}$ define a given binary characteristic of the individual measured in 1995 (such as his/her gender or whether he/she works as an employee or a self-employed). For every possible dimension of heterogeneity, I estimate the following fully interacted versions of specifications (2.10) and (2.11):[28]

$$y_{itd} = \lambda_i + \gamma_{td} + \sum_{k=t_0}^{t_1} \beta_k T_i \cdot \mathbb{1}(\text{year} = k) + \sum_{k=t_0}^{t_1} \beta_k^d T_i \cdot \mathbb{1}(\text{year} = k) \cdot d + \varepsilon_{it} \qquad (2.12)$$

and

$$y_{itd} = \lambda_i + \gamma_{td} + \beta T_i \cdot Post_t + \beta^d T_i \cdot Post_t \cdot d + \varepsilon_{it} \qquad (2.13)$$

2.6 Results

2.6.1 Main effects

Figure 2.5 displays estimates based on specification (2.10).[29] The five outcomes are: the probability that at least one day in a given year is covered by SS contributions (panel a); the number of days per year covered by SS contributions (panel b); the probability that at least one day in a given year is covered by work-related SS contributions (panel c); the number of days per year covered by work-related SS contributions (panel d); yearly labour earnings (panel e). In 4 out of 5 cases pre-1995 coefficients, which are mostly positive, are small and not significantly different from 0: this suggests that *treated* and *control* individuals were on substantially parallel trends before the *Dini* reform. In panel (a) one of the pre-1995 coefficients is significantly different from 0 at the 5 per cent level and three are at the 10 per cent level: in this case there appear to exist small deviations from parallel trends in the period preceding the *Dini* reform. Focusing on post-1995 coefficients, in the first four panels they are positive and significantly different from 0 starting from 1997 or 1998; moreover, coefficients increase over time, although the growth flattens out in the final years of the sample period. This pattern suggests that there is a positive effect on labour supply. The size of the effect is small, resulting in a narrower gap in labour supply between *treated* and *control* individuals than that documented in

[27] This choice of the window allows one to maintain a large enough sample size while excluding individuals very close to the threshold.

[28] These fully interacted specifications deliver results that are the same as those that would be obtained by running separate regressions on the relevant sub-samples; they make it easier to assess whether the effects found in the various sub-groups of the population are significantly different.

[29] Coefficients are also reported in Appendix Table 2.E.1.

1995 (Table 2.1): with regards to the number of days covered by work-related SS contributions, for instance, the estimated coefficient is equal to almost 3 in 1998 and to 5.5 in 2006, against an average of 292 days per year among *treated* individuals in the pre-1996 period. Focusing on labor earnings, standard errors are bigger, but post-reform coefficients also in this case are positive, increasing over time and significantly different from 0 at the 5 or 10 per cent level in most years.[30]

Table 2.2 reports results from the more compact specification (2.11): the probability of having at least one day in a given year covered by SS contributions or by work-related SS contributions increases on average in the post-1995 period by 0.6 and 0.9 percentage points, respectively. The coefficient relative to the number of days covered by SS contributions is 3.3, amounting to 1.08 per cent of the pre-1996 average value among *treated* individuals; the same figure for the number of days covered by work-related SS contributions is 3.8 (1.28 per cent). Finally, labour earnings increase by 198 euros per year, although the coefficient is not statistically significant at conventional levels.

Overall, these results suggest that changes to the way of calculating Social Security benefits may induce small labour supply responses from individuals who are still relatively far away from retirement benefit eligibility. The fact that post-1995 coefficients increase over time could reflect a number of reasons. First, there could be adjustments costs of a various nature that make changes to labour supply happen gradually. Second, rules for computing retirement benefits are complex: it may, therefore, also takes time to fully understand the changes brought to them. At the same time, as individuals get older, the way in which their public pension is calculated may become more salient. Third, over 1996-2006 the 5-year moving average of the real GDP growth rate peaked at the beginning of the 2000s and then started to decline. Some *treated* individuals may have recognized that this would impact negatively the NDC portion of their pension benefits.

2.6.2 Robustness checks

First, I check whether results change when including observations very close to the 18-year threshold that separates *treated* individuals from *control* ones. Figure 2.6 and Table 2.3 present results based on specifications (2.10) and (2.11), respectively.[31] They show that results are virtually unchanged when including these individuals.[32]

Second, I perform two placebo exercises. In the first one, I pretend that the *Dini* reform had set the threshold that separates fully grandfathered individuals from partially grandfathered ones at 19 rather than at 18 years of SS contributions: *placebo*

[30] As explained in Section 2.4, the number of observations in the regression for labour earnings is lower due to some missing values.

[31] Coefficients displayed in Figure 2.6 are also reported in Appendix Table 2.E.2.

[32] Pre-1995 coefficients in panel (a) are slighly farther away from 0.

treated are individuals with 18 to 19 (excluded) years of SS contributions, whereas *placebo controls* are workers with 19 (included) to 20 years of SS contributions. In the second one, I assume that the threshold had been set at 20 years of contributions: *placebo treated* are then individuals with 19 to 20 (excluded) years of SS contributions, whereas *placebo controls* are workers with 20 (included) to 21 years of SS contributions. I continue restricting the attention to individuals aged 35-45 in 1995 and excluding observations in the 4-week window on either side of the placebo thresholds. All individuals in the placebo samples are not affected by changes to the way of calculating retirement benefits brought about by the *Dini* reform.

Figures 2.7 and 2.8 display the results of these exercises. Focusing on outcomes in panels (a) to (d), in both figures post-1995 coefficients are smaller than those coming from the estimates based on the true 18-year threshold and in virtually all cases they are not statistically significant.[33] Also estimates based on the more compact specification (2.11) deliver coefficients that are smaller and non statistically significant (Table 2.4). However, there are two notes of caution. First, in the exercise that sets the placebo threshold at 19 years of contributions β_{1990} and β_{1991} are significantly different from 0, weakening the assumption that *placebo treated* and *placebo controls* were on parallel trends since the beginning of the sample period before the *Dini* reform. Second, in the exercise that sets the placebo threshold at 20 years of contributions, coefficients relative to earnings (panel e) are in some cases not much smaller than the true ones. Overall, placebo exercises suggest that the effects documented in Subsection 2.6.1 are at least partially reflecting responses to the changes in the computation of pension benefits, and not only differential trends in labour supply among individuals with slightly different contribution histories up to 1995. At the same time, some caution in the interpretation of results is required.

2.6.3 Heterogeneous effects

Having discussed the main findings in Subsection 2.6.1 and a battery of robustness checks in Subsection 2.6.2, in this subsection I perform a series of heterogeneity analyses. First, I consider each possible dimension of heterogeneity separately, by estimating specifications (2.12) and (2.13); then, using an enriched version of specification (2.13), I estimate an "horserace" regression whereby I consider all dimensions of heterogeneity together.

I start by checking whether the strength of labour supply responses varies by gender. Figure 2.9 shows that the adjustments along the margin of days covered by (all and work-related only) SS contributions documented in the whole sample seem to entirely reflect the behaviour of women, while men do not seem to adjust their labour supply in a significant way.[34] Post-1995 female coefficients are significantly larger

[33] Coefficients are also reported in Appendix Tables 2.E.3 and 2.E.4.

[34] Coefficients displayed in Figure 2.9 are also reported in Appendix Table 2.E.5.

than male ones: as an example, with regards to days covered by work-related SS contributions, the coefficient in 2006 is 10.9 for women, while it is less than one third of this (2.8) for men. This also clearly emerges in Table 2.5, which displays results based on the more compact specification (2.11). With respect to labour earnings, the picture is more nuanced and coefficients are more similar to each other; column 5 of Table 2.5 shows that the percentage increase is larger for women, who earned less before 1996.[35]

Among the cohorts of workers studied, the older ones may pay more attention to changes of the public pension system. Furthermore, older individuals have less time left to adjust their labour supply relatively to younger ones. For all these reasons, I study whether labour supply responses differ across narrower age groups: specifically, I separate the youngest cohorts - *i.e.* those aged 35-37 - from the other ones. Figure 2.10 displays the results.[36] With respect to days covered by SS contributions, the post-1995 coefficients of older cohorts are very similar to those of younger ones in all but the last two years, when the difference becomes positive. However, the coefficients relative to the years 1990-1992 of the younger cohorts are significantly different from 0, suggesting a deviation from parallel trends. For this reason, the positive differential effect found in Table 2.6, which is based on the more compact pre-post 1995 comparison, should be interpreted with caution. When focusing on days covered by work-related SS contributions, the pre-1995 coefficients of younger cohorts are closer to 0 and the post-1995 coefficients are lower than those of older cohorts, although the difference is not statistically significant. Also in this case, the positive differential effect that emerges from Table 2.6 should, therefore, be interpreted with caution. Furthermore, when looking at labour earnings, it seems that the response is concentrated among younger individuals. Overall, this heterogeneity exercise provides inconclusive results.

As shown in Section 2.3, grandfathering clauses embedded in the 1992 and 1995 pension reforms created a complex web of rules for computing retirement benefits. It would therefore be interesting to study whether labour supply adjustments vary depending on the level of education and financial literacy. For instance, Banks et al. (2015) document that financial literacy is an important determinant of annuitization choices over private DC pension wealth made by older English individuals. Unfortunately, administrative data does not record information about financial literacy or the highest educational attainment. I build a proxy of education based on the age at which an individual has the first year-round (i.e. covering 52 weeks) contribution spell. If this happens before age 25, I classify an individual as having

[35]However, such difference has to be interpreted with caution as it also reflects a different behaviour of earnings among men and women in the pre-reform period: while female coefficients are negative, although not significantly different from 0, male coefficients are positive and significantly different from 0 in 1993 and 1994.

[36]Coefficients are also reported in Appendix Table 2.E.6.

a low or medium level of education; if this happens between age 25 and age 32, I define an individual as having a high level of education.[37] This measure of education is a crude proxy and is subject to classification error: however, individuals labeled as highly educated earn more than others in the pre-1996 period, which would be consistent with the existence of a positive return to education. Another inevitable drawback is that, given the cohorts studied, more educated individuals are also on average older: hence, it is difficult to separate the effect of education from the effect of age. This is one of the reasons why at the end of this subsection I also estimate a regression whereby I consider all dimensions of heterogeneity together. Figure 2.11 display estimates based on specification (2.12).[38] Labour supply responses seem to be of similar magnitude in the immediate aftermath of the *Dini* reform. Starting from year 2001 they, however, become larger among highly-educated individuals and the difference is statistically significant in the final years of the sample period. As an example, with regards to the number of days covered by SS contributions, the coefficient in 2006 is 12.8 for highly-educated individuals and 4.4 for other workers. This significant heterogeneity also emerges from Table 2.7, which reports results based on the more compact specification (2.13).

Due to the fact that the yield rate declines with earnings in the Quota B formula, having part of retirement benefits computed according to the NDC method is likely to penalize less or benefit high earners (see Section 2.3). It is, therefore, interesting to divide individuals based on the earnings they were making in the pre-1996 period: I create two groups depending on whether workers belong to the bottom and middle tercile or to the top tercile of the distribution.[39] Figure 2.12 displays estimates based on specification (2.12).[40] It shows that labour supply responses documented in the whole sample seem to reflect the behaviour of individuals belonging to the bottom and middle terciles of the earnings distribution: neither the number of days nor labour earnings of high earners change significantly in the post-1996 period.[41] These findings are confirmed in Table 2.8, which reports estimates based on the more compact specification (2.13).

As explained in Section 2.3, the yield rate in the NDC system is smaller for the self-employed than for employees, due to the lower social security tax rate. As a result, the former are affected to a larger extent than the latter from having part

[37] I do not classify individuals who have the first full year covered by SS contributions after age 32. This is the reason why this set of regressions features fewer observations.

[38] These coefficients are also reported in Appendix Table 2.E.7.

[39] Specifically, individuals are ranked according to their labour earnings in year z, the last pre-reform year in which they had worked at least one day; these earnings are adjusted by expressing them in 1995 euros and by applying a penalty factor that is quadratic in the distance between year z and 1995, to take into account that periods of non-employment often cause skill depreciation.

[40] These coefficients are also reported in Appendix Table 2.E.8.

[41] With respect to the number of days covered by SS contributions, the difference between high earners and others is significant in most post-reform years; with respect to earnings, coefficients are similar but those of high earners are not statistically different from 0.

of retirement benefits computed according to the NDC formula. I therefore study whether labour supply adjustments are different across salaried workers and independent ones. With regards to the number of days covered by (all and work-related only) SS contributions, Figure 2.13 shows that there are no substantial differences in the post-1995 period; moreover, pre-1995 coefficients are instead in some instances different across employees and self-employed, so that the overall results look fuzzy. Nonetheless, earnings of the self-employed increase by more. The difference, although substantial, only becomes significant in the very final years of the sample period, partly due to the larger standard errors surrounding the estimates for the self-employed.[42] The same pattern emerges also from Table 2.9. While no conclusive answer can be given, it could be that the tighter link between benefits and contributions incentives the self-employed, who could more easily avoid declaring some income, to report more labour earnings.

So far, I have investigated each possible dimension of heterogeneity separately. However, differences along one dimension could be correlated with differences in other ones. As an example women, who have been found to be more responsive than men, are also slightly older and have lower earnings in the pre-1996 period. I, therefore, conclude this subsection by considering all dimensions of heterogeneity together, based on an enriched version of specification (2.13), whereby I interact the term $T \cdot Post$ and year fixed effects with a set of dummy variables capturing all such dimensions. Table 2.10 shows the results of this exercise. In general, coefficients retain the same sign, but decrease in magnitude. The difference between females and males remains significant at the 10 per cent level with respect to the probability of having at least one day covered by work-related contributions and the number of such days. The same holds true for the difference in terms of earnings between employees and the self-employed. Apart from those, the only coefficients that preserve their significance are the ones related to age - suggesting a stronger response of older cohorts - when looking at the number of days covered by (all and work-related only) days of contributions. However, given the issues described above with the exercise that divides individuals by age, I also present a version of this exercise whereby I exclude age as an interaction term. Appendix Table 2.E.10 shows that, in this case, differences between women and men and between low-medium and high educated individuals in terms of days covered by SS contributions would remain significant.

2.7 Conclusions

Spurred by the pressures that population ageing exerts on the sustainability of public pension systems, many countries have implemented reforms that reduce the generosity of retirement benefits and tighten eligibility criteria. While the ultimate

[42] These coefficients are also reported in Appendix Table 2.E.9.

goal of these reforms is to foster labour market participation at older ages, they are often announced years before affected individuals come close to the decision of when to retire and claim benefits. Due to grandfathering clauses and gradual phase-in schedules, new provisions indeed often do not apply to workers on the cusp of retirement, whilst partially or fully affecting younger ones. It is therefore interesting to study not only the effects of pension reforms on labour supply at mature ages, but also since the time of the announcement, as well as to asses how adjustments unfold over time.

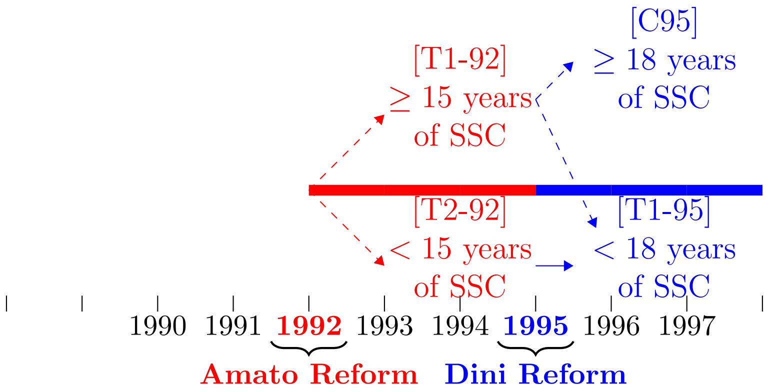
In this paper I have attempted to address these questions by observing over the period 1990-2006 the labour supply trajectories of a sample of Italian private-sector workers who were aged 35-45 in 1995, when a pension reform changed the way in which a part of their retirement benefits would be calculated. The reform was likely to lower the level of retirement benefits at early benefit claiming ages. Moreover, it increased the difficulty of forecasting such level, by making it subject to macro-economic and demographic risks. Furthermore, it tightened the link between contributions paid to Social Security and benefits received, on top of increasing incentives to delay claiming benefits past eligibility.

I have leveraged administrative data on full contribution histories and, within a difference-in-differences framework, I have compared these individuals with same-age ones who were entirely grandfathered because they happened to have slightly more contributions to Social Security by the end of 1995. While estimates have to be taken with some caution, I found over the period of the analysis modest positive effects on labour supply of affected workers, as measured by days covered by (all and work-related only) contributions to Social Security and by labour earnings. The effects tend to grow over time. This pattern may stem from a variety of reasons that is not possible to disentangle with the available data. It could be due to adjustment costs embedded in labour supply decisions; it could also reflect a process of gradual learning, whereby individuals become more aware about their retirement prospects and the features of the public pension system as they age.

It would be interesting to see if these findings generalize to other reform episodes and countries. Furthermore, in my data I cannot observe hours worked, which are however an important margin of labour supply; having data on them available in other settings would shed more light on the anatomy of labour supply adjustments.

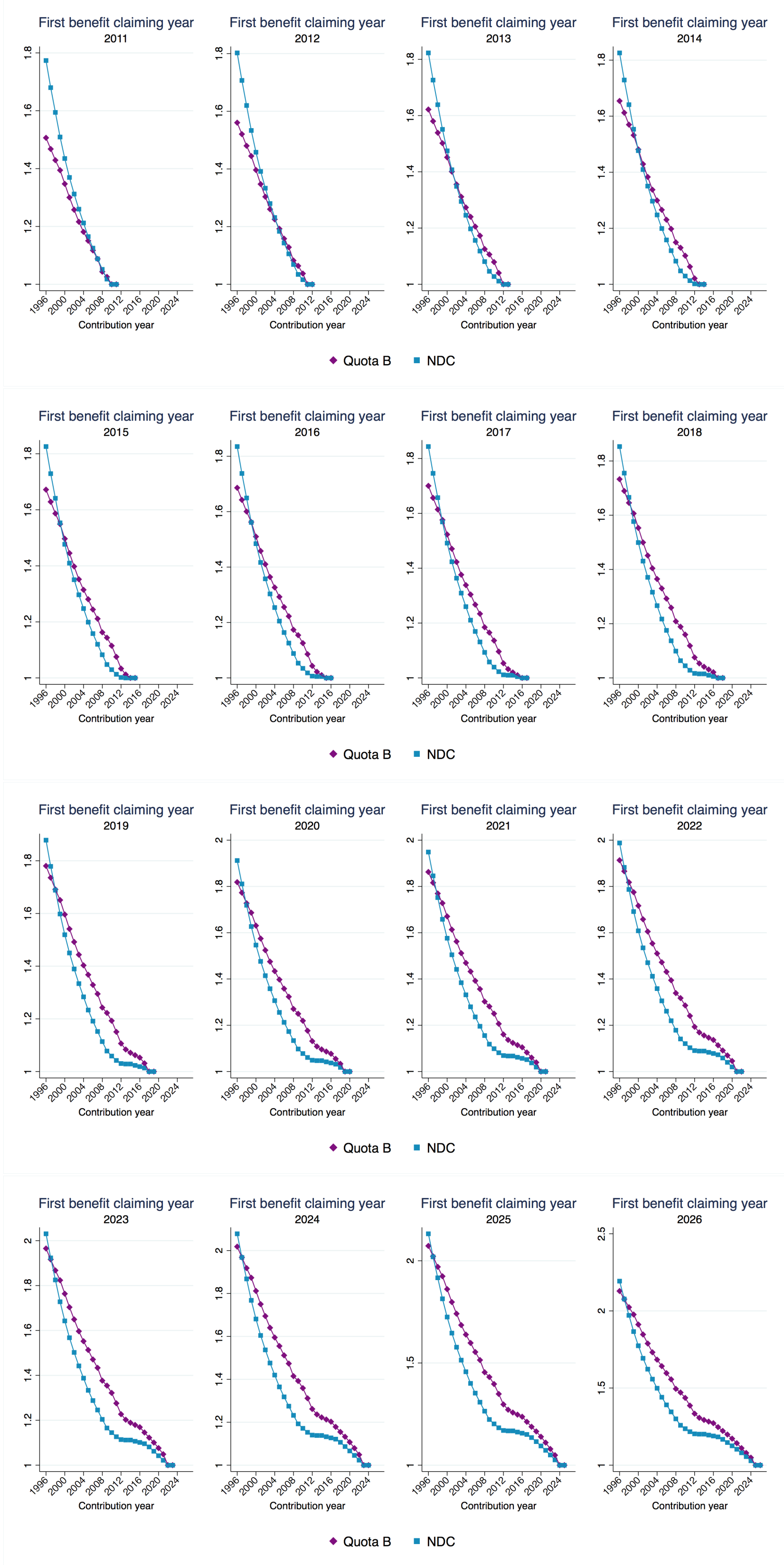
2.8 Figures

Figure 2.1: The 1992 *Amato* and the 1995 *Dini* pension reforms



Notes: This figure reports the timeline of the 1992 *Amato* and the 1995 *Dini* reform, as well as the different groups created by grandfathering clauses.

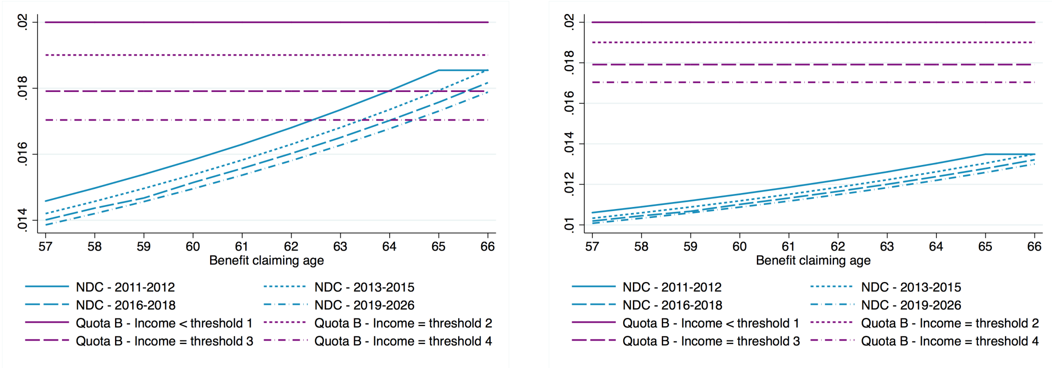
Figure 2.2: Accrual rate in the Quota B and in the NDC regimes



Notes: The figure plots the accrual rate in the Quota B and in the NDC regimes, by first benefit claiming year.

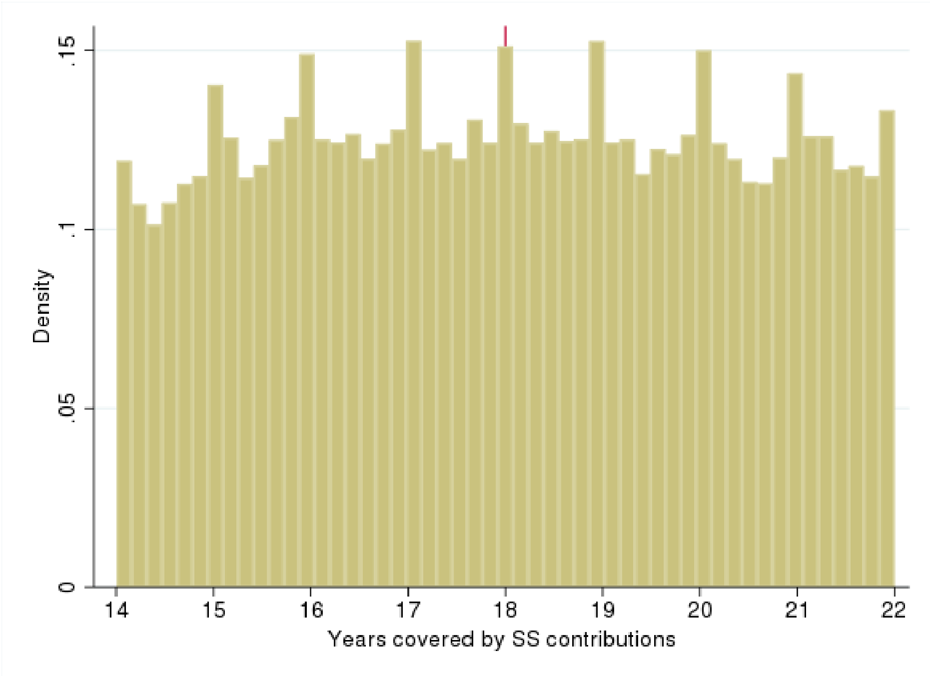
Figure 2.3: Yield rate in the Quota B and in the NDC regimes

(a) Employees (b) Self-employed



Notes: Panels (a) and (b) plot the yield rate in the Quota B and in the NDC regimes for employees and for the self-employed, respectively. In the Quota B regime the yield rate does not vary with the age at retirement, but it starts at 2 per cent and then declines as average earnings grow and surpass various thresholds. In the figure I plot the yield rate for average earnings below the first threshold, as well as the average yield rate for average earnings that equal exactly the following thresholds. In the NDC system the yield rate is the product between the social security tax rate, if constant, and the transformation coefficient (i.e. the coefficient that transforms the notional capital into an annuity). The social security tax rate is 33 per cent for employees and has grown from 20 to 24 per cent for the self-employed: in the figure and in the formulae in the text, the Social Security tax rate for the self-employed is assumed constant and set to 24 per cent. This explains why the yield rate is larger for the former than for the latter. The transformation coefficient is an increasing function of the age at retirement, thus making the yield rate higher for individuals who start claiming retirement benefits later. The transformation coefficient was revised downward, to account for increases in life expectancy, in 2010, 2013, 2016 and 2019.

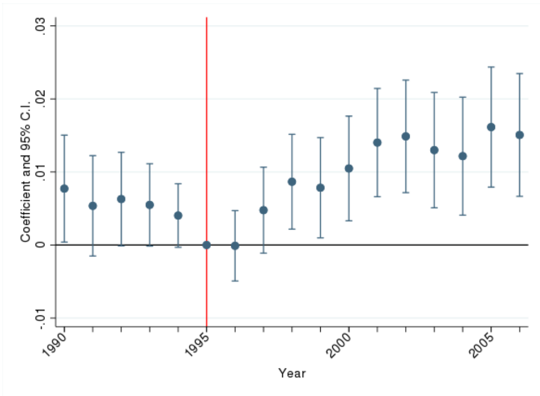
Figure 2.4: Distribution of years of Social Security contributions by 1995 for workers aged 35-45 and with 14 to 22 years of contributions



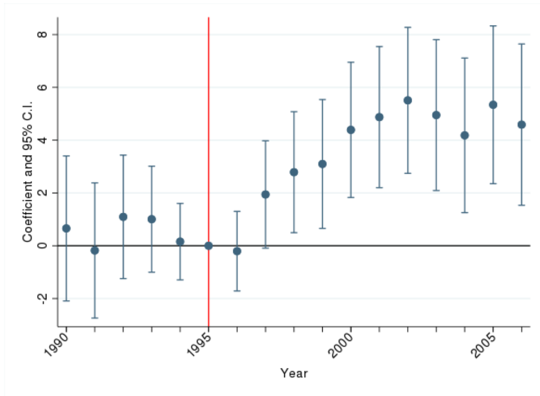
Notes: The figure plots the density histogram of years of Social Security contributions by the end of 1995. The sample consists of individuals aged 35-45 in 1995 who had accumulated by that time 14 to 22 years of SS contributions.

Figure 2.5: Main results

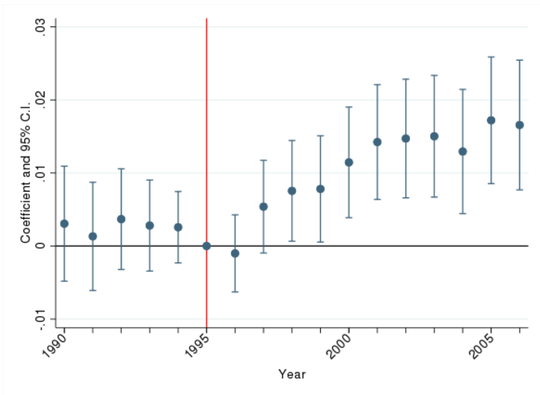
(a) Contributions to SS (1=yes)



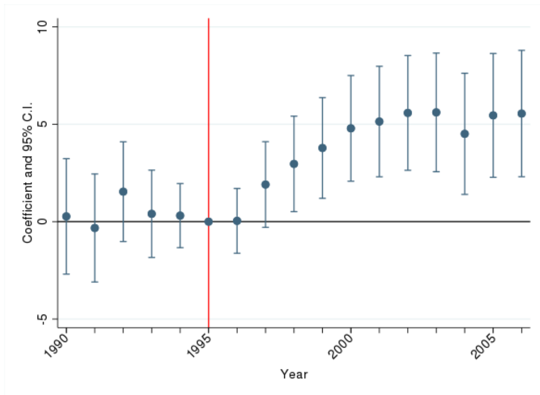
(b) Contributions to SS (days)



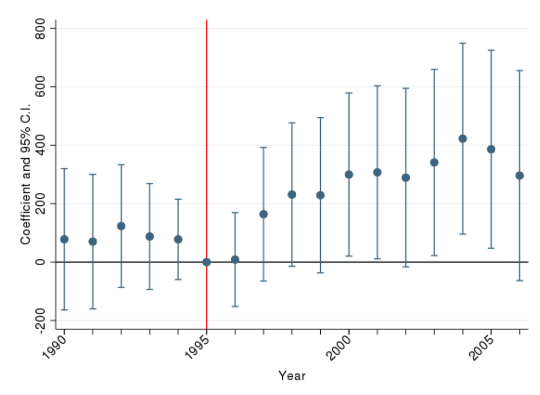
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



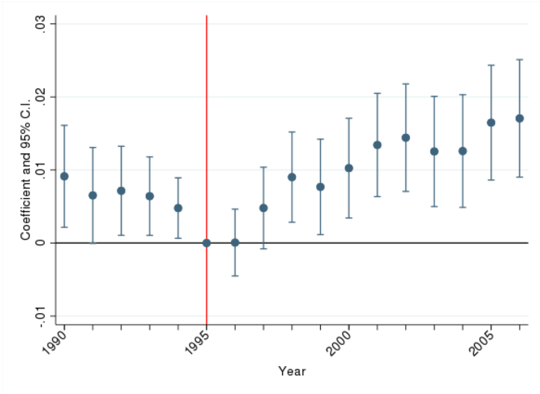
(e) Yearly labour earnings (euros)



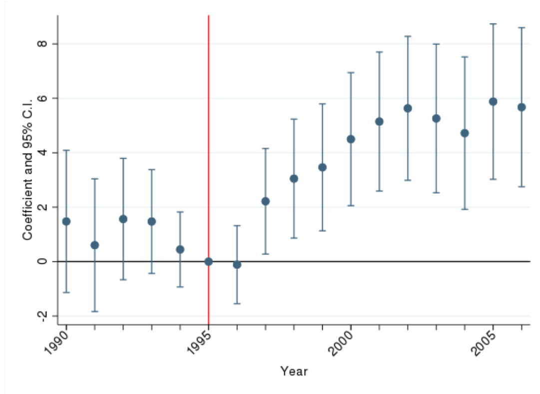
Notes: Panels (a) to (e) plot coefficients β_ks from specification (2.10), which includes individual and year fixed effects. Dots show the difference between *treated* and *control* workers in year k relative to the year when the *Dini* reform was passed (1995). Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. Standard errors are clustered at the individual level.

Figure 2.6: Main results, including observations very close to the threshold

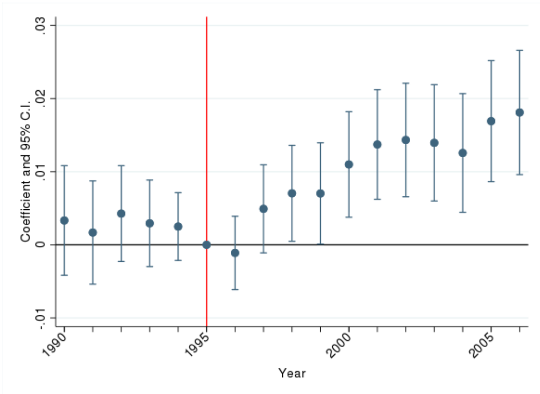
(a) Contributions to SS (1=yes)



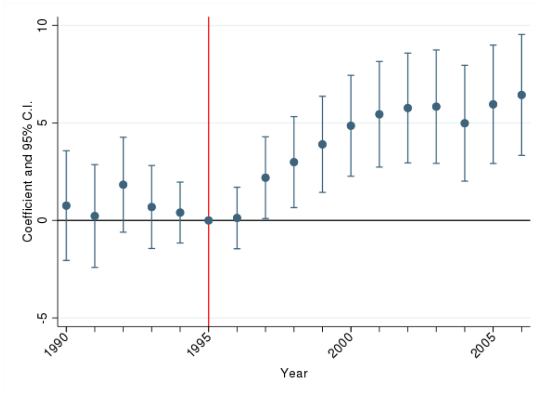
(b) Contributions to SS (days)



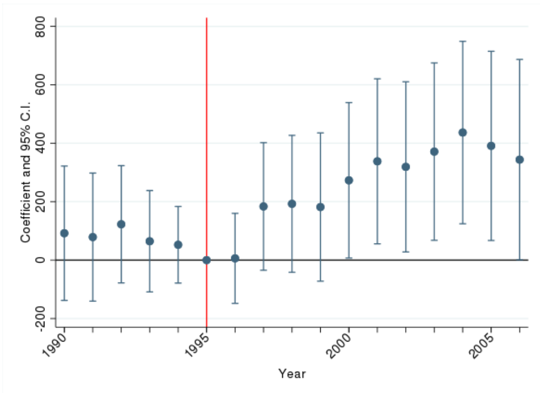
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



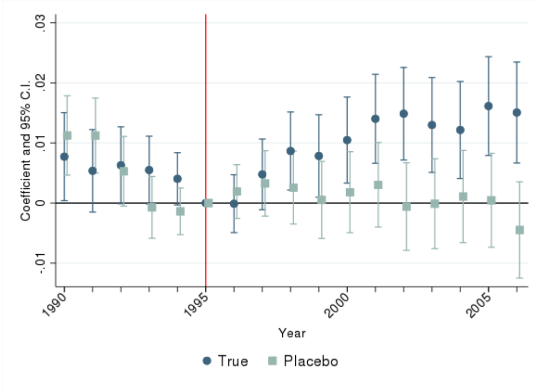
(e) Yearly labour earnings (euros)



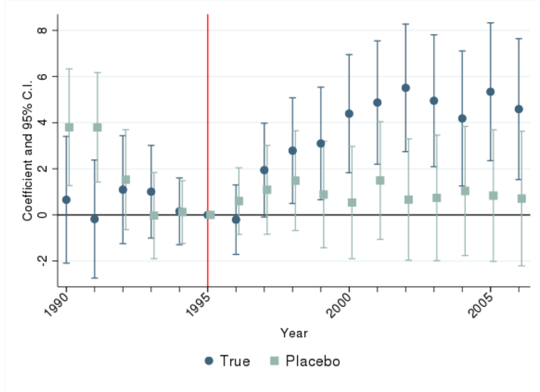
Notes: Panels (a) to (e) plot coefficients β_ks from specification (2.10), which includes individual and year fixed effects. Dots show the difference between *treated* and *control* workers in year k relative to the year when the *Dini* reform was passed (1995). Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Standard errors are clustered at the individual level.

Figure 2.7: Placebo threshold at 19 years of SS contributions

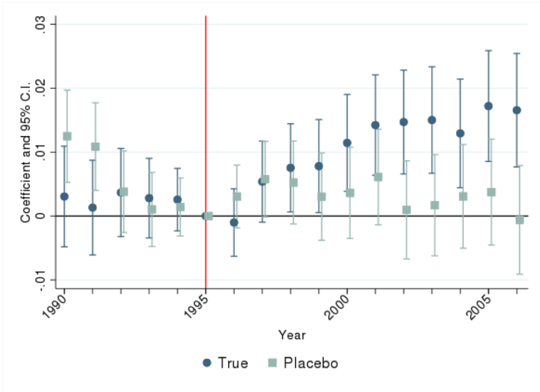
(a) Contributions to SS (1=yes)



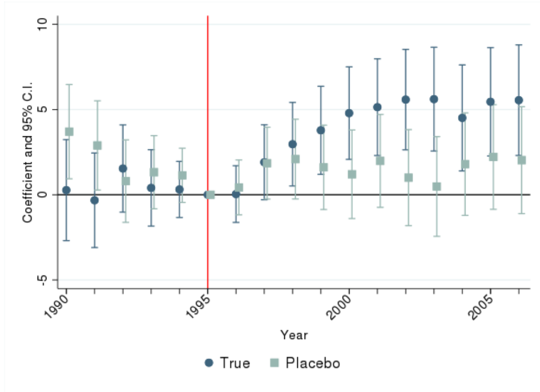
(b) Contributions to SS (days)



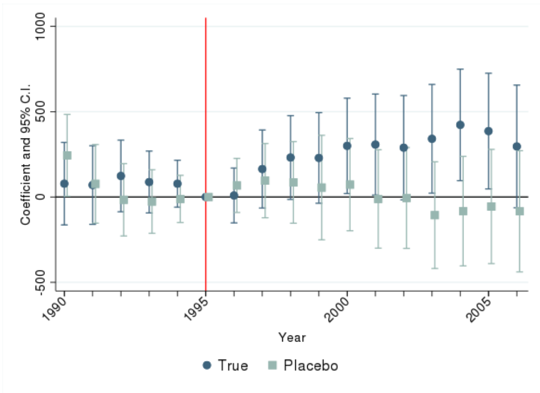
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



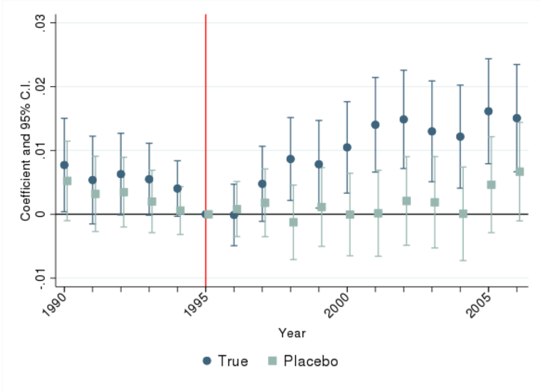
(e) Yearly labour earnings (euros)



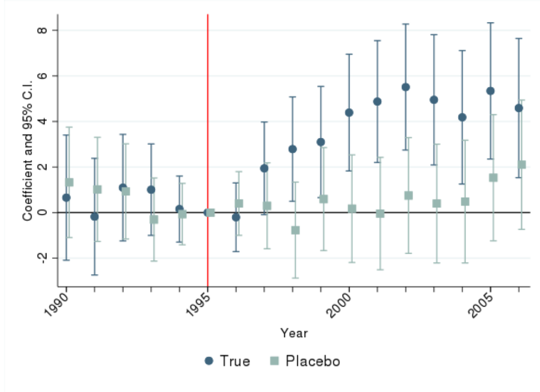
Notes: The blue dots show the difference between *true treated* and *true control* individuals in year k relative to the year when the *Dini* reform was passed (1995). The light green squares capture the difference between *placebo treated* and *placebo control* individuals when the placebo threshold is set at 19 years of contributions. The definition of *placebo treated* and *placebo controls* is provided in Subsection 2.6.2. Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Estimates are based on the difference-in-differences specification (2.10), which includes individual and year fixed effects. Observations lying in the 4-week window on either side of the 18-year or the 19-year thresholds are excluded. Standard errors are clustered at the individual level.

Figure 2.8: Placebo threshold at 20 years of SS contributions

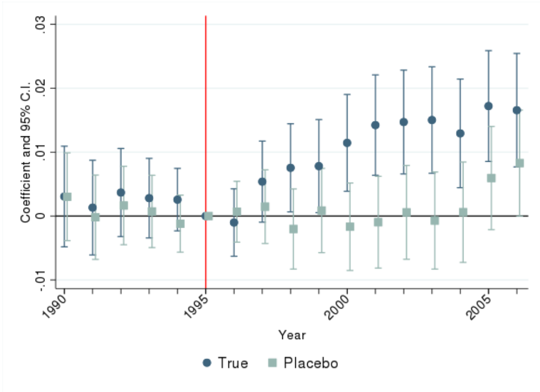
(a) Contributions to SS (1=yes)



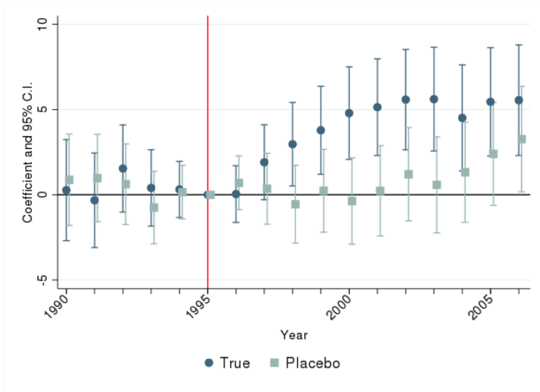
(b) Contributions to SS (days)



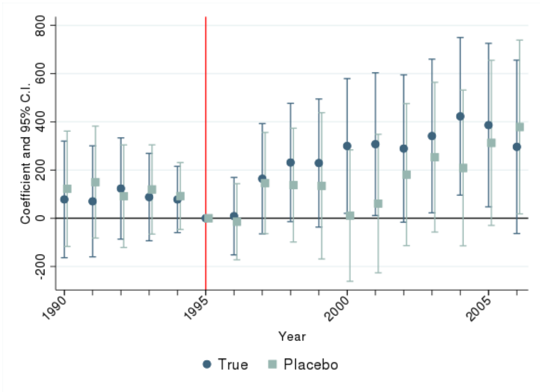
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



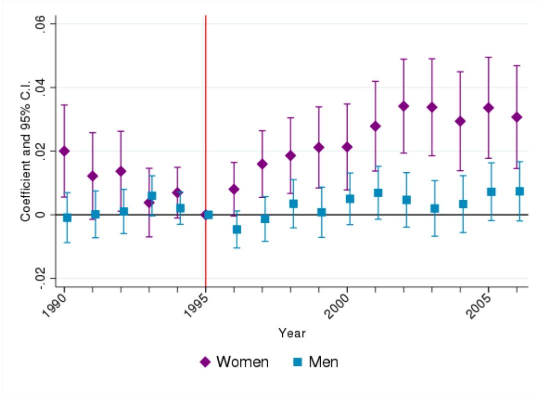
(e) Yearly labour earnings (euros)



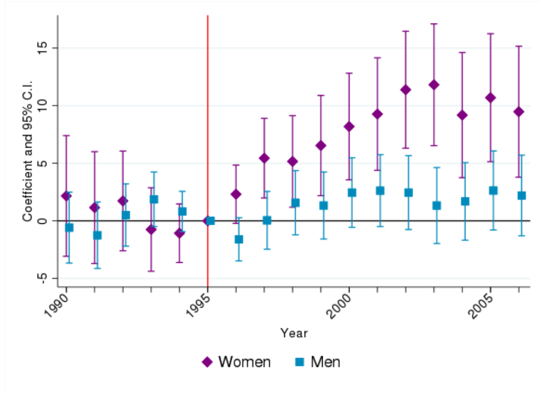
Notes: The blue dots show the difference between *true treated* and *true control* individuals in year k relative to the year when the *Dini* reform was passed (1995). The light green squares capture the difference between *placebo treated* and *placebo control* individuals when the placebo threshold is set at 20 years of contributions. The definition of *placebo treated* and *placebo controls* is provided in Subsection 2.6.2. Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Estimates are based on the difference-in-differences specification (2.10), which includes individual and year fixed effects. Observations lying in the 4-week window on either side of the 18-year or the 20-year thresholds are excluded. Standard errors are clustered at the individual level.

Figure 2.9: Main results, heterogeneity by gender

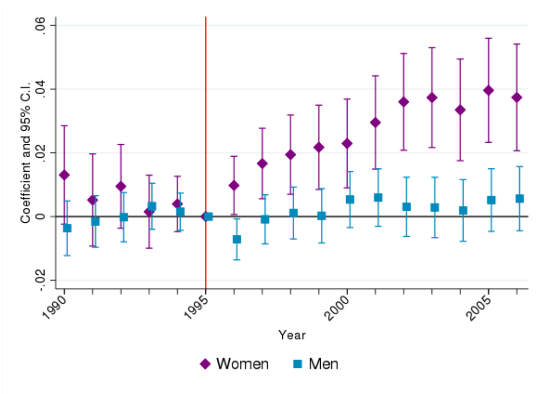
(a) Contributions to SS (1=yes)



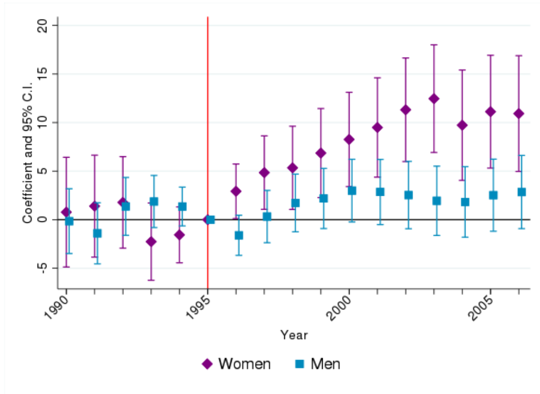
(b) Contributions to SS (days)



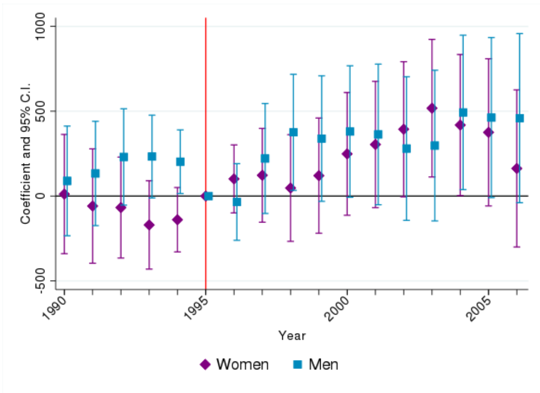
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



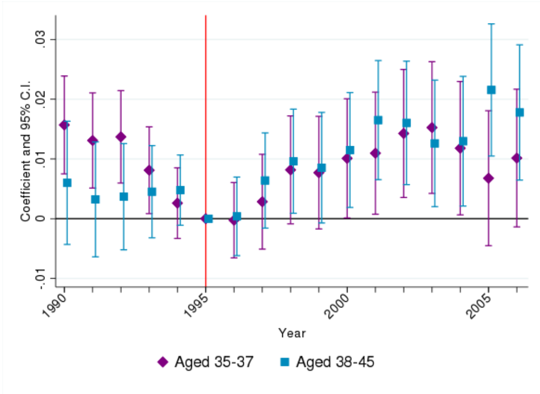
(e) Yearly labour earnings (euros)



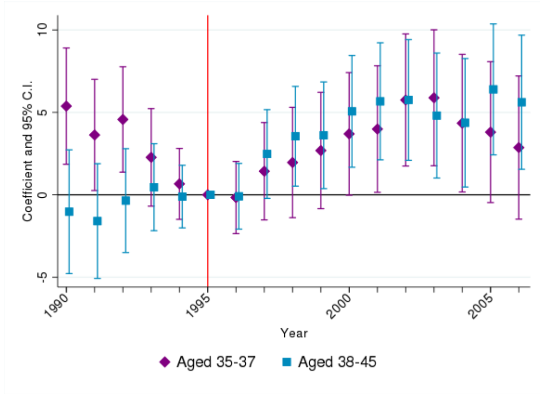
Notes: Panels (a) to (e) plot coefficients β_ks and $(\beta_k + \beta_k^d)$s from specification (2.12), which includes individual and year-gender fixed effects. Diamonds and squares show the difference between *treated* and *control* workers in year k relative to the year when the *Dini* reform was passed (1995). Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. Standard errors are clustered at the individual level.

Figure 2.10: Main results, heterogeneity by age in 1995

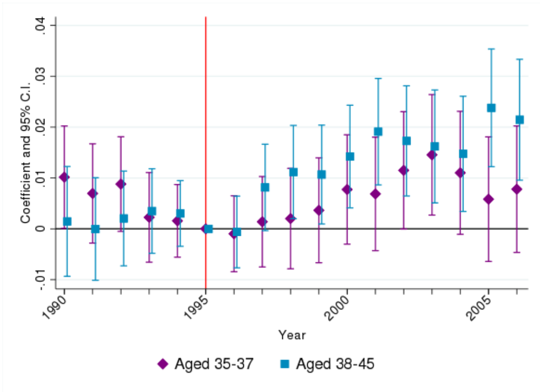
(a) Contributions to SS (1=yes)



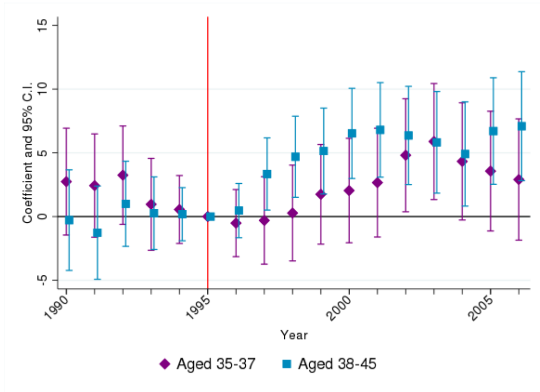
(b) Contributions to SS (days)



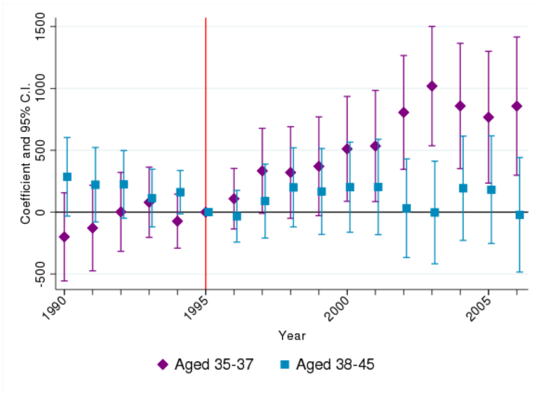
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



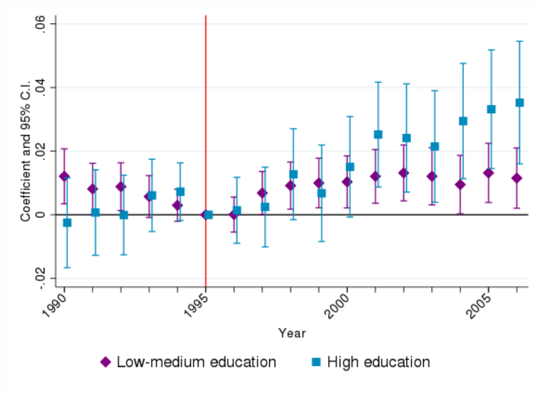
(e) Yearly labour earnings (euros)



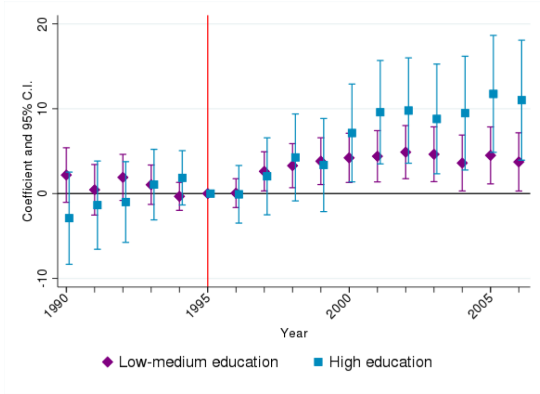
Notes: Panels (a) to (e) plot coefficients β_ks and $(\beta_k + \beta_k^d)$s from specification (2.12), which includes individual and year-age group fixed effects. Diamonds and squares show the difference between *treated* and *control* workers in year k relative to the year when the *Dini* reform was passed (1995). Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. Standard errors are clustered at the individual level.

Figure 2.11: Main results, heterogeneity by (a proxy of) education

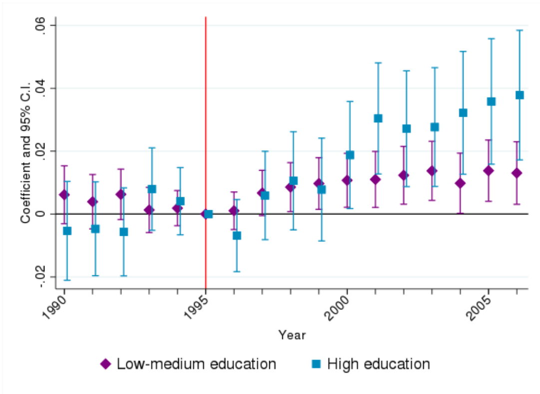
(a) Contributions to SS (1=yes)



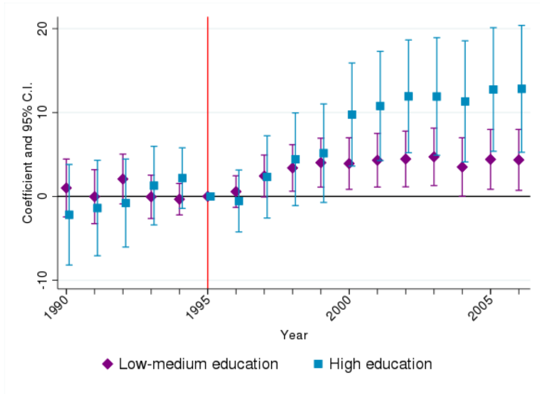
(b) Contributions to SS (days)



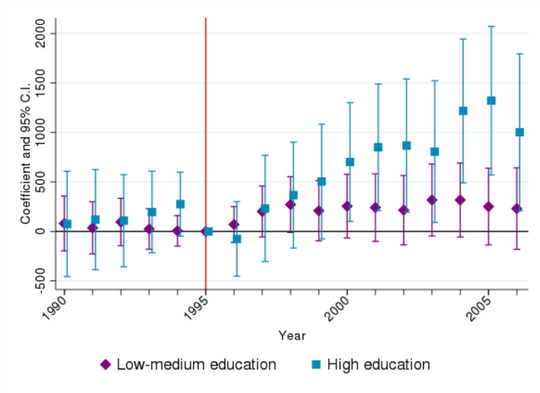
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



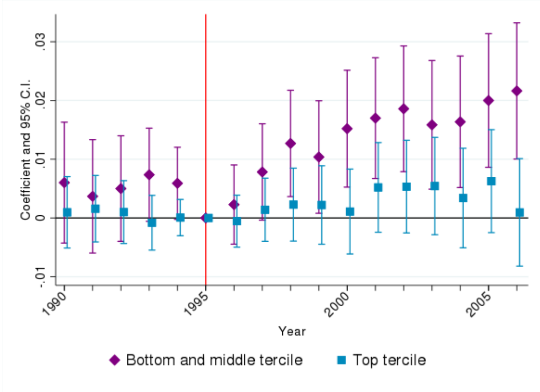
(e) Yearly labour earnings (euros)



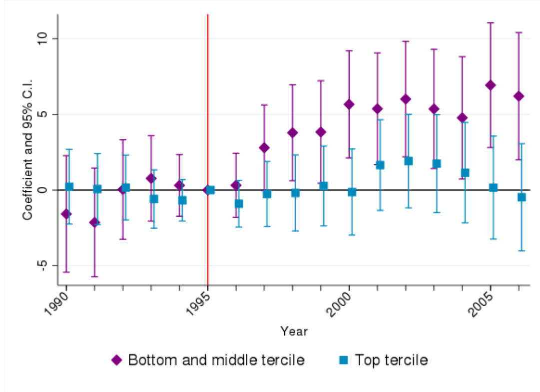
Notes: Panels (a) to (e) plot coefficients β_ks and $(\beta_k + \beta_k^d)$s from specification (2.12), which includes individual and year-education group fixed effects. Diamonds and squares show the difference between *treated* and *control* workers in year k relative to the year when the *Dini* reform was passed (1995). Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. Standard errors are clustered at the individual level.

Figure 2.12: Main results, heterogeneity by pre-1996 labour earnings

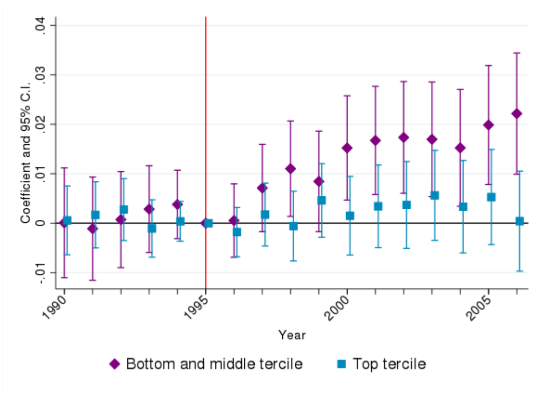
(a) Contributions to SS (1=yes)



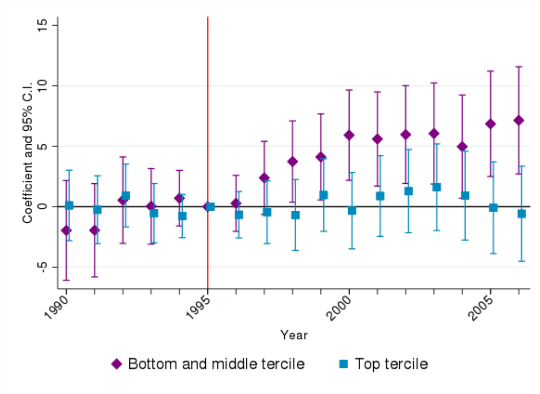
(b) Contributions to SS (days)



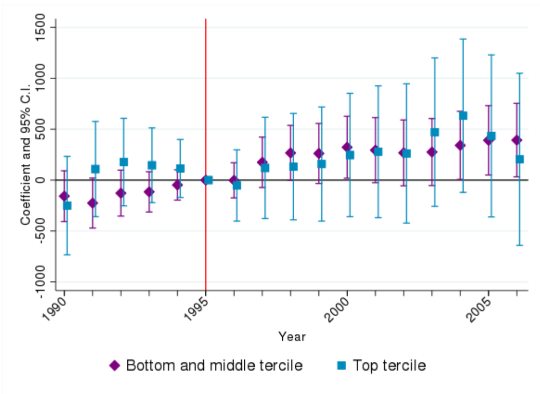
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



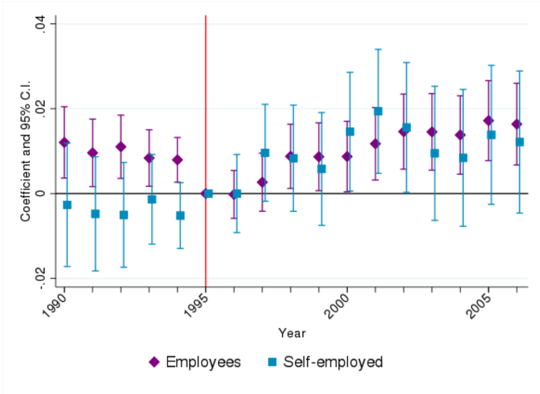
(e) Yearly labour earnings (euros)



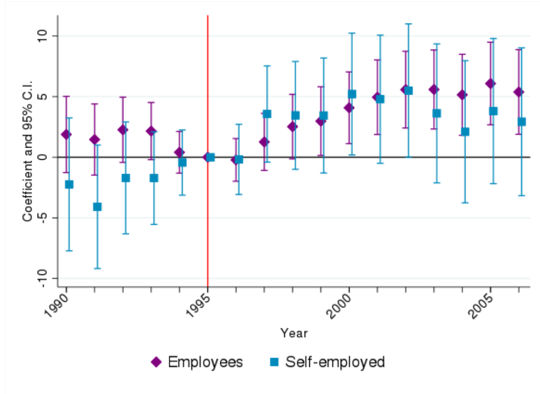
Notes: Panels (a) to (e) plot coefficients β_ks and $(\beta_k + \beta_k^d)$s from specification (2.12), which includes individual and year-income group fixed effects. Diamonds and squares show the difference between *treated* and *control* workers in year k relative to the year when the *Dini* reform was passed (1995). Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. Standard errors are clustered at the individual level.

Figure 2.13: Main results, heterogeneity by pre-1996 occupation

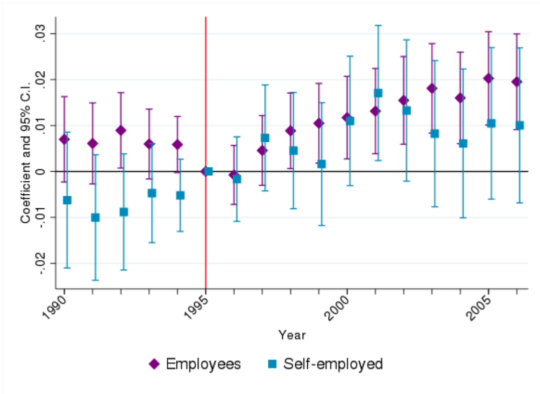
(a) Contributions to SS (1=yes)



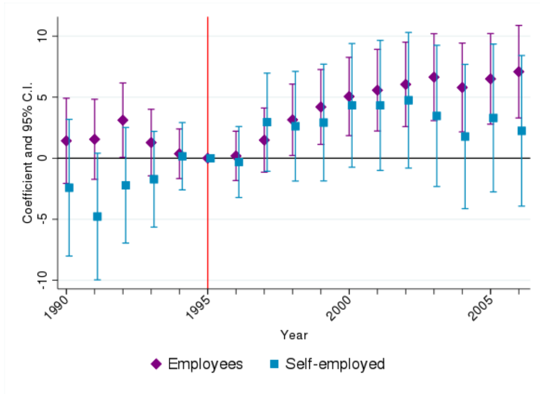
(b) Contributions to SS (days)



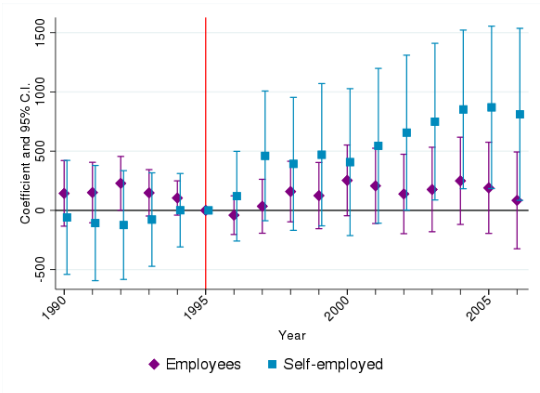
(c) Work-related contributions to SS (1=yes)



(d) Work-related contributions to SS (days)



(e) Yearly labour earnings (euros)



Notes: Panels (a) to (e) plot coefficients β_ks and $(\beta_k + \beta_k^d)$s from specification (2.12), which includes individual and year-occupation fixed effects. Diamonds and squares show the difference between *treated* and *control* workers in year k relative to the year when the *Dini* reform was passed (1995). Whiskers represent 95 per cent confidence intervals. The outcomes are: (a) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (b) the number of days per year covered by SS contributions; (c) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (d) the number of days per year covered by work-related SS contributions; (e) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. Standard errors are clustered at the individual level.

2.9 Tables

Table 2.1: Sample summary statistics

	Control		Treated	
	mean	sd	mean	sd
	(1)	(2)	(3)	(4)
Gender (1=male)	0.64	0.48	0.62	0.48
Age in 1995	39.25	2.97	39.15	2.99
Place of birth (1= Italy)	0.98	0.13	0.98	0.14
Occupation (1=self-employed)	0.31	0.46	0.30	0.46
Work-related SSCs (1=yes)	0.86	0.34	0.83	0.38
Work-related SSCs (days)	296.33	132.64	280.97	143.29
SSCs (1=yes)	0.89	0.32	0.85	0.35
SSCs (days)	308.23	121.49	293.89	133.67
Labour earnings (euros)	20649.84	15292.91	19679.73	15700.65
Observations	17955		17188	

Notes: The table reports summary statistics as of 1995 for individuals belonging to the *treated* and the *control* groups. The information about the place of birth is missing for around 15 per cent of individuals and the share of those born in Italy is computed over non-missing observations.

Table 2.2: Main results

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.006**	3.313***	0.009***	3.755***	197.672
	(0.003)	(1.052)	(0.003)	(1.101)	(123.104)
Observations	543507	543507	543507	543507	538773
R-squared	0.586	0.627	0.592	0.628	0.767
$\bar{y}^T_{\text{pre}}$	0.90	307.39	0.86	292.31	19990.01
Percentage effect	0.65	1.08	1.02	1.28	0.99

Notes: The table reports estimates of the coefficient β from specification (2.11), which includes individual and year fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. In the row "Percentage effect" coefficients are divided by the average value of the outcome among *treated* workers in the pre-1996 period ($\bar{y}^T_{\text{pre}}$). Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.3: Main results, including observations very close to the threshold

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.005*	3.203***	0.008***	3.755***	207.721*
	(0.003)	(1.004)	(0.003)	(1.050)	(117.549)
Observations	597431	597431	597431	597431	592095
R-squared	0.585	0.626	0.591	0.626	0.768
$\bar{y}^T_{\text{pre}}$	0.90	308.15	0.87	293.30	20115.30
Percentage effect	0.57	1.04	0.96	1.28	1.03

Notes: The table reports estimates of the coefficient β from specification (2.11), which includes individual and year fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. In the row "Percentage effect" coefficients are divided by the average value of the outcome among *treated* workers in the pre-1996 period ($\bar{y}^T_{\text{pre}}$). Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.4: Placebo thresholds at 19 and 20 years of SS contributions

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
Panel a: placebo threshold at 19 years of contributions					
T · Post	-0.003	-0.621	-0.002	-0.124	-41.792
	(0.003)	(1.009)	(0.003)	(1.062)	(121.628)
Observations	540056	540056	540056	540056	535730
R-squared	0.573	0.614	0.581	0.615	0.763
$\bar{y}^T_{\text{pre}}$	0.92	320.27	0.89	306.51	21013.52
Percentage effect	-0.37	-0.19	-0.19	-0.04	-0.20
Panel b: placebo threshold at 20 years of contributions					
T · Post	-0.001	0.054	0.001	0.538	68.651
	(0.003)	(0.991)	(0.003)	(1.045)	(122.292)
Observations	521815	521815	521815	521815	517568
R-squared	0.563	0.602	0.569	0.603	0.764
$\bar{y}^T_{\text{pre}}$	0.94	329.18	0.92	315.88	21994.53
Percentage effect	-0.08	0.02	0.06	0.17	0.31

Notes: The table reports estimates of the coefficient β from specification (2.11), which includes individual and year fixed effects. In panel (a) the placebo threshold of the *Dini* reform that separates fully grandfathered from partially grandfathered individuals is set at 19 years of contributions; in panel (b) it is set at 20 years of contributions. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. In the row "Percentage effect" coefficients are divided by the average value of the outcome among *placebo treated* workers in the pre-1996 period ($\bar{y}^T_{\text{pre}}$). Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.5: Main results, heterogeneity by gender

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.016***	7.598***	0.022***	8.458***	326.337**
	(0.006)	(2.027)	(0.006)	(2.099)	(161.911)
T · Post · Male	-0.014**	-6.303***	-0.020***	-7.134***	-143.707
	(0.006)	(2.334)	(0.007)	(2.434)	(234.177)
Observations	543507	543507	543507	543507	538773
R-squared	0.589	0.629	0.594	0.628	0.768
$\bar{y}^{T,F}_{\text{pre}}$	0.83	275.99	0.77	255.84	14596.88
$\bar{y}^{T,M}_{\text{pre}}$	0.94	326.46	0.92	314.45	23239.41
Percentage effect F	1.88	2.75	2.86	3.31	2.24
Percentage effect M	0.19	0.40	0.24	0.42	0.79

Notes: The table reports estimates of the coefficients β and β^d from specification (2.13), which includes individual and year-gender fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. In the rows "Percentage effect F" and "Percentage effect M" coefficients are divided by the average value of the outcome among *treated* female and male workers in the pre-1996 period ($\bar{y}^{T,F}_{\text{pre}}$ and $\bar{y}^{T,M}_{\text{pre}}$). Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.6: Main results, heterogeneity by age in 1995

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.000	0.542	0.002	0.836	643.157***
	(0.004)	(1.486)	(0.004)	(1.584)	(186.101)
T · Post · Age 38-45	0.008	4.188**	0.011*	4.440**	-700.767***
	(0.005)	(2.045)	(0.006)	(2.155)	(245.638)
Observations	543507	543507	543507	543507	538773
R-squared	0.586	0.628	0.592	0.628	0.767
$\bar{y}_{\text{pre}}^{T,35-37}$	0.98	342.13	0.95	329.62	22381.98
$\bar{y}_{\text{pre}}^{T,38-45}$	0.85	287.67	0.81	271.13	18617.32
Percentage effect 35-37	0.00	0.16	0.16	0.25	2.87
Percentage effect 38-45	0.99	1.64	1.54	1.95	-0.31

Notes: The table reports estimates of the coefficients β and β^d from specification (2.13), which includes individual and year-age group fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. In the rows "Percentage effect 35-37" and "Percentage effect 38-45" coefficients are divided by the average value of the outcome among *treated* younger and older workers in the pre-1996 period ($\bar{y}_{\text{pre}}^{T,35-37}$ and $\bar{y}_{\text{pre}}^{T,38-45}$). Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.7: Main results, heterogeneity by (a proxy of) education

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.004	2.735**	0.007**	3.201**	194.399
	(0.003)	(1.193)	(0.003)	(1.245)	(141.013)
T · Post · High edu	0.013*	4.667*	0.014**	5.370*	385.088
	(0.007)	(2.640)	(0.007)	(2.775)	(304.831)
Observations	532746	532746	532746	532746	528251
R-squared	0.589	0.630	0.595	0.629	0.767
$\bar{y}^{pre}_{T,\text{low-medium}}$	0.88	302.22	0.85	287.86	19598.89
$\bar{y}^{pre}_{T,\text{high}}$	0.93	323.91	0.90	308.99	21746.60
Percentage effect low-med. edu.	0.40	0.91	0.80	1.11	0.99
Percentage effect high edu.	1.81	2.29	2.36	2.77	2.66

Notes: The table reports estimates of the coefficients β and β^d from specification (2.13), which includes individual and year-education group fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. In the rows "Percentage effect low-med. edu." and "Percentage effect high edu." coefficients are divided by the average value of the outcome among *treated* low-medium and high educated workers in the pre-1996 period ($\bar{y}^{pre}_{T,\text{low-medium}}$ and $\bar{y}^{pre}_{T,\text{high}}$). Standard errors are reported in parenthesis and are clustered at the individual level. *, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.8: Main results, heterogeneity by pre-1996 labour earnings

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.010**	5.075***	0.012***	5.261***	383.812***
	(0.004)	(1.451)	(0.004)	(1.513)	(127.355)
T · Post · Top tercile	-0.007	-4.491**	-0.010**	-4.911**	-170.636
	(0.005)	(1.834)	(0.005)	(1.964)	(300.208)
Observations	531845	531845	531845	531845	530008
R-squared	0.583	0.620	0.584	0.615	0.768
$\bar{y}_{\text{pre}}^{T,\text{bottom and middle}}$	0.87	292.22	0.83	274.12	14067.25
$\bar{y}_{\text{pre}}^{T,\text{top}}$	0.98	353.97	0.98	350.12	34303.34
Percentage effect bot./mid.	1.12	1.74	1.50	1.92	2.73
Percentage effect top	0.26	0.16	0.22	0.10	0.62

Notes: The table reports estimates of the coefficients β and β^d from specification (2.13), which includes individual and year-income group fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. In the rows "Percentage effect bot./mid." and "Percentage effect top" coefficients are divided by the average value of the outcome in the pre-1996 period among *treated* workers in the bottom and middle terciles or in the top tercile of the earnings distribution ($\bar{y}_{\text{pre}}^{T,\text{bottom and middle}}$ and $\bar{y}_{\text{pre}}^{T,\text{top}}$). Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.9: Main results, heterogeneity by pre-1996 occupation

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.002	2.581**	0.007**	3.412***	14.249
	(0.003)	(1.206)	(0.003)	(1.274)	(140.129)
T · Post · Self-emp.	0.011*	2.594	0.007	1.358	622.626**
	(0.006)	(2.408)	(0.007)	(2.485)	(285.006)
Observations	543507	543507	543507	543507	538773
R-squared	0.586	0.628	0.593	0.629	0.767
$\bar{y}_{\text{pre}}^{T,\text{employee}}$	0.90	304.08	0.85	283.45	21478.64
$\bar{y}_{\text{pre}}^{T,\text{self-emp}}$	0.89	315.10	0.89	312.95	16531.74
Percentage effect employee	0.27	0.85	0.80	1.20	0.07
Percentage effect self-emp.	1.54	1.64	1.56	1.52	3.85

Notes: The table reports estimates of the coefficients β and β^d from specification (2.13), which includes individual and year-occupation fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. In the rows "Percentage effect employees" and "Percentage effect self-emp." coefficients are divided by the average value of the outcome among *treated* employees and the self-employed in the pre-1996 period ($\bar{y}_{\text{pre}}^{T,\text{employee}}$ and $\bar{y}_{\text{pre}}^{T,\text{self-emp}}$). Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.10: Main results, heterogeneity along all dimensions together

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.003	3.650	0.009	4.658*	481.999**
	(0.007)	(2.565)	(0.007)	(2.702)	(241.884)
T · Post · Male	-0.009	-3.979	-0.012*	-4.268*	-22.240
	(0.007)	(2.442)	(0.007)	(2.563)	(245.619)
T · Post · Age 38-45	0.011*	4.927**	0.013**	4.665*	-424.610
	(0.006)	(2.283)	(0.006)	(2.408)	(270.850)
T · Post · High edu.	0.009	2.589	0.010	3.534	455.439
	(0.008)	(2.896)	(0.008)	(3.051)	(334.738)
T · Post · Self emp.	0.009	1.381	0.005	0.118	567.492*
	(0.007)	(2.549)	(0.007)	(2.647)	(302.021)
T · Post · Top terc.	-0.002	-2.408	-0.005	-3.029	-18.856
	(0.005)	(2.008)	(0.006)	(2.163)	(324.787)
Observations	521611	521611	521611	521611	519824
R-squared	0.588	0.625	0.589	0.619	0.770

Notes: The table reports estimates based on an enriched version of specification (2.13), where the term $T_i \cdot Post_t$ and year fixed effects are also interacted with: a dummy that takes value 1 if the individual is male; a dummy that takes value 1 if the individual is aged 38-45 in 1995; a dummy that takes value 1 if the individual is high-educated (as measured by a proxy); a dummy that takes value 1 if the individual belongs to the top tercile of the distribution of pre-1996 labour earnings; a dummy that takes value 1 if the individual worked as a self-employed in the last occupation before 1996. Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Appendices

2.A Additional details on the retirement benefit formulae

Following the 1992 *Amato* reform, first year annual pension benefits were computed as the sum of two "quotas", Quota A and Quota B:

$$b = \underbrace{k_A \cdot N_A \cdot \bar{y}_{t_l:R}}_{\text{Quota A}} + \underbrace{k_B \cdot N_B \cdot \bar{y}_{t_q:R}}_{\text{Quota B}} \tag{2.14}$$

N_A were years of SS contributions up to 1992, while N_B were years of SS contributions from 1993 to retirement R. If the sum of N_A and N_B exceeded 40, the excess years beyond the fortieth were not taken into account for computing the "quota" that featured the lowest average earnings. Therefore:

$$\begin{aligned} N_A = {} & N_{t_0:1992} \cdot \mathbb{1}(N_{t_0:R} <= 40 \vee \bar{y}_{t_l:R} >= \bar{y}_{t_q:R}) + \\ & + (40 - N_{1993:R}) \cdot \mathbb{1}(N_{t_0:R} > 40 \wedge \bar{y}_{t_l:R} < \bar{y}_{t_q:R}) \end{aligned} \tag{2.15}$$

$$\begin{aligned} N_B = {} & N_{1993:R} \cdot \mathbb{1}(N_{t_0:R} <= 40 \vee \bar{y}_{t_q:R} >= \bar{y}_{t_l:R}) + \\ & + (40 - N_{t_0:1992}) \cdot \mathbb{1}(N_{t_0:R} > 40 \wedge \bar{y}_{t_q:R} < \bar{y}_{t_l:R}) \end{aligned} \tag{2.16}$$

The accrual rate in the Quota B part of (2.14) was higher than the accrual rate in the Quota A part: beyond taking into account inflation, it added a fixed 1 per cent premium for every year until retirement. Therefore:

$$\Psi_{B,t} = \prod_{j=t}^{R} (1 + \pi_j + 1\%) \text{ for } t = t_q, ..., R-2 \text{ and } \Psi_{B,R-1} = \Psi_{B,R} = 1 \tag{2.17}$$

The *Amato* reform aimed at gradually lengthening the reference period over which to compute average earnings. Specifically, for individuals with at least 15 years of SS contributions by 1992 (T1-92), the reference period of Quota B was extended to the last M weeks covered by SS contributions, being M equal to 520 for employees and to 780 for self-employed. For individuals with less than 15 years of SS contributions by 1992 (T2-92), it included all O weeks covered by SS contributions between 1993 and retirement, and the last P weeks covered by SS contributions before 1993, being P equal to 260 for employees and to 520 for self-employed. Finally, for individuals who started contributing in 1993 or after (T3-92), the reference period was the entire contribution life. Therefore, for a given individual i:

$$\bar{y}_{t_q:R} = \begin{cases} \bar{y}_{t_v:R} = \frac{1}{M/52}\sum_{t=t_v}^{R} y_t\Psi_{B,t}, & \text{if } i \in \text{T1-92} \\ \bar{y}_{t_w:R} = \frac{1}{((O+P)/52)}\sum_{t=t_w}^{R} y_t\Psi_{B,t}, & \text{if } i \in \text{T2-92} \\ \bar{y}_{t_0:R} = \frac{1}{Q/52}\sum_{t=t_0}^{R} y_t\Psi_{B,t}, & \text{if } i \in \text{T3-92} \end{cases} \tag{2.18}$$

with:

$$t_v : \sum_{j=t_v+1}^{R} \frac{n_j}{52} < \frac{M}{52} \leq \sum_{j=t_v}^{R} \frac{n_j}{52} \text{ and } t_w : \sum_{j=t_w+1}^{R} \frac{n_j}{52} < \frac{O+P}{52} \leq \sum_{j=t_w}^{R} \frac{n_j}{52} \tag{2.19}$$

Q represents the number of weeks covered by SS contributions from the beginning of the contribution history to its end.

2.B Simulation of the change of first year pension benefits

I simulate the expected change, as of 1995, of first year pension benefits due to the *Dini* reform. I focus on individuals belonging to the T1-95 group and I restrict the attention to the portion of first year pension benefits associated to post-1995 contributions to Social Security. I then proceed according to the following steps:

1. I assume that daily earnings (y^d) and the number of days worked (d) in a given post-1995 year t are given by the following formulae:

$$y_t^d = y_0^d \prod_{j=1996}^{j=t} (1 + \pi_j + \mu) \text{ where } y_0^d = y_z^d \big(1 + \frac{z - 1995}{100}\big)^2 \tag{2.20}$$

$$d_t = k \tag{2.21}$$

z is the last pre-1996 year such that at least x days are covered by work-related contributions to Social Security, with $x = \{1{,}90{,}180\}$. The starting daily earnings y_0^d are daily earnings made in year z, expressed in 1995 euros and adjusted for a penalty that is quadratic in the distance between year z and 1995, to account for the fact that periods of suspended or reduced activity in the labour market can cause skill depreciation and thus lead to a lower labour income when an individual starts working again. Daily earnings y_0^d then grow every year at a rate that equals the sum of the realized inflation π and a parameter μ that captures real earnings growth. In alternative parametrizations, μ can take the following values: 0 per cent; 1 per cent; 1.5 per cent; 2 per cent; 2.5 per cent; 3 per cent.

Individuals are assumed to work a fixed number of days per year (k) in the post-1995 period. k can take two values: first, $k = d_z$, i.e. individuals work the same number of days as they did in year z; alternatively, it is assumed that all individuals hold year-round jobs in the post-reform period ($d = f$).

2. I assume that individuals start collecting retirement benefits from January, 1st of year $R + 1$. I make two alternative assumptions about R, which is the last contribution year. First, I assume it is the earliest year in which individuals become eligible to claim either an old age or a seniority pension under the rules prevailing in 1995. Hence $R = min(t_o, t_s)$, where t_o is the year when women (resp. men) turn 60 (resp. 65) and t_s is the first year when individuals have both at least 57 years of age and at least 35 years of contributions, or 40 years of contributions regardless of age. Second, I assume that workers stop contributing once they become eligible to claim an old-age pension, i.e. $R = t_o$. Given that the master sample consists of workers aged 35 to 45 in

1995, R spans from 2010 to 2025.

3. I assume that every worker will contribute in the post-1995 period into the same pension fund (i.e. the fund of employees or the funds of self-employed) as he/she did in year z. This assumption is needed because different funds feature different Social Security tax rates τ. I am therefore implicitly assuming no switches from salaried employment to self-employment, or the other way around, in the post-1995 period.

4. I assume that workers foresee the evolution over time of the parameters embedded in the Quota B and NDC formulae of yearly pension benefits. I therefore assume that $\{\Psi_B,\ k_B\}$ and $\{G,\ \tau,\ \beta_a\}$ take the realized values. For cohorts of workers who retire past 2019, I predict the values that these parameters will take by relying on available forecasts of the inflation rate and the nominal GDP growth rate.

5. Given $\{y_t^d, d_t, \pi_t, \mu, \Psi_{B,t}, k_{B,t}\}_{t=1996}^{t=R}$, I compute first year annual pension benefits stemming from post-1995 contributions to Social Security that workers belonging to the T1-95 group would have been entitled to, had the *Dini* reform not passed. In this counterfactual scenario, post-1995 contributions would concur to determine first year pension benefits according to the following formula:

$$\bar{b}_{\text{post-95}} = k_B \cdot N_{1996:R} \cdot \bar{y}_{t_q:R} \tag{2.22}$$

   where the definition of $\bar{y}_{t_q:R}$ is provided in Appendix Section 2.A.[43]

6. Given $\{y_t^d, d_t, \pi_t, \mu, G_t, \tau_t, \beta_{a,t}\ \}_{t=1996}^{t=R}$, I compute first year annual pension benefits stemming from post-1995 contributions to Social Security that a worker belonging to the T1-95 group is entitled to after the *Dini* reform. Post-1995 contributions concur to determine first year pension benefits according to the following NDC formula:

$$\hat{b}_{\text{post-95}} = \beta_a \cdot \left(\sum_{t=1996}^{R} \tau_t y_t G_t \right) \tag{2.23}$$

7. I express the change θ in the portion of first year pension benefits stemming from post-1995 contributions to Social Security in the following way:

$$\theta = 100 \cdot \frac{\hat{b}_{\text{post-95}} - \bar{b}_{\text{post-95}}}{\bar{b}_{\text{post-95}}} \tag{2.24}$$

[43] For individuals with less than 15 years of contributions by the end of 1992, computing $\bar{y}_{t_q:R}$ also requires to know earnings between 1993 and 1995 as well as in the last 5 years of contributions to Social Security before 1993 (10 years for self-employed). I retrieve this information from observed pre-1996 contribution histories.

Table 2.B.1 shows the results of the simulations under the different parametrizations of μ, d and R. x is set equal to 1.[44] It shows the average change (column 1), as well changes corresponding to the 25th, 50th and 75th percentiles of the distribution (column 2-4); it also reports the share of individuals who would experience a negative change (column 5). The average and the median changes are negative across all simulations; moreover the percentage of individuals who would receive lower benefits is always higher than 70 per cent. The loss is lower under the assumption that individuals wait to claim retirement benefits until they become eligible to obtain an old-age pension. This is due to three main reasons. First, by retiring weakly later than under the alternative assumption, pensionable earnings are higher and thus are more likely to surpass the threshold above which the average yield rate in the Quota B system declines. This is also the reason why the loss is lower at higher values of the growth rate of real earnings. Second, retiring weakly later implies a higher probability of hitting and surpassing the 40-year threshold after which years of SS contributions are not taken into account in the Quota B formula.[45] Third, the 5-year moving average of the nominal GDP growth rate, which defines the accrual rate in the NDC system, has been particularly low in the mid-2010s, following the double dip recession that hit Italy, while it is expected to increase in the early 2020s as the economy recovers.

[44] Simulations under the alternative values of x deliver very similar results.

[45] This happens if average earnings in the Quota B formula are lower than average earnings in the Quota A formula.

Table 2.B.1: Simulation of the change in first year annual pension benefits stemming from post-1995 contributions

	mean (1)	p25 (2)	p50 (3)	p75 (4)	% neg. (5)
x =1, μ =0, d =f, R = $min(t_o, t_s)$	-36.30	-50.80	-32.34	-26.86	1.00
x =1, μ =0, d =f, R = t_o	-29.79	-45.32	-24.06	-19.92	0.99
x =1, μ =0, d =d_z, R = $min(t_o, t_s)$	-36.40	-51.02	-32.68	-26.79	0.99
x =1, μ =0, d =d_z, R = t_o	-30.80	-46.21	-25.01	-20.43	0.99
x =1, μ =1, d =f, R = $min(t_o, t_s)$	-35.46	-50.07	-31.07	-26.49	1.00
x =1, μ =1, d =f, R = t_o	-26.03	-39.44	-22.93	-18.10	0.91
x =1, μ =1, d =d_z, R = $min(t_o, t_s)$	-34.94	-49.79	-30.98	-25.76	0.99
x =1, μ =1, d =d_z, R = t_o	-26.33	-42.37	-22.95	-17.94	0.91
x =1, μ =1.5, d =f, R = $min(t_o, t_s)$	-35.01	-49.67	-30.69	-26.12	1.00
x =1, μ =1.5, d =f, R = t_o	-21.61	-33.66	-22.82	-5.54	0.82
x =1, μ =1.5, d =d_z, R = $min(t_o, t_s)$	-34.18	-49.10	-30.28	-25.11	0.99
x =1, μ =1.5, d =d_z, R = t_o	-21.60	-35.13	-22.31	-5.55	0.82
x =1, μ =2.0, d =f, R = $min(t_o, t_s)$	-34.53	-49.21	-30.06	-25.65	0.99
x =1, μ =2.0, d =f, R = t_o	-19.52	-32.02	-20.51	-3.21	0.79
x =1, μ =2.0, d =d_z, R = $min(t_o, t_s)$	-33.40	-48.37	-29.53	-24.40	0.99
x =1, μ =2.0, d =d_z, R = t_o	-19.70	-33.31	-20.74	-3.49	0.79
x =1, μ =2.5, d =f, R = $min(t_o, t_s)$	-34.02	-48.44	-29.47	-24.96	0.99
x =1, μ =2.5, d =f, R = t_o	-18.58	-32.36	-20.63	-1.75	0.77
x =1, μ =2.5, d =d_z, R = $min(t_o, t_s)$	-32.60	-47.66	-28.74	-23.55	0.98
x =1, μ =2.5, d =d_z, R = t_o	-18.39	-32.55	-20.16	-1.73	0.77
x =1, μ =3.0, d =f, R = $min(t_o, t_s)$	-33.49	-47.59	-28.92	-24.14	0.99
x =1, μ =3.0, d =f, R = t_o	-17.40	-32.92	-20.60	-0.11	0.75
x =1, μ =3.0, d =d_z, R = $min(t_o, t_s)$	-31.77	-46.85	-27.95	-22.64	0.98
x =1, μ =3.0, d =d_z, R = t_o	-16.96	-32.20	-19.52	0.02	0.75

Notes: The table reports the results of simulations of the expected change in first year annual pension benefits stemming from post-1995 contributions, under alternative parametrization of μ, d and R. x is set equal to 1.

2.C Computing years of SS contributions by the end of 1995

To identify workers belonging to the C95 and T1-95 groups, I need to compute weeks of qualifying contributions to Social Security by the end of 1995. Given that the unit of observation of the main dataset is the individual-contribution spell-year, I first need to compute the number of weeks of qualifying contributions in any given year before 1996 and then sum them up. To this end, I follow strictly the dispositions provided by the Italian Institute of Social Security (INPS), whereby in case of overlapping contribution spells in any given year only one of them shall be considered useful toward the acquisition of pension rights and overlapping contributions shall not be summed.

Let d^s_{ijy} be the recorded start date of a given spell i of individual j in year y and d^e_{ijy} be the recorded end date; c_{ijy} are the recorded weeks of contributions associated to that spell. In some cases, the number of days between the recorded start and end dates of a given spell is different from the recorded days of contributions to Social Security associated to such a spell. To take into account the rule about overlapping spells, I proceed in the following way:

1. For every contribution spell $[d^s_{ijy}, d^e_{ijy}]$ I compute the average number of contribution weeks accrued in a given day d_{ijy} belonging to such a spell ($\bar{c}^d_{ijy}$), that is:
$$\forall d_{ijy} \in [d^s_{ijy}, d^e_{ijy}], \quad \bar{c}^d_{ijy} = \frac{c_{ijy}}{d^e_{ijy} - d^s_{ijy}} \tag{2.25}$$

2. If a given day d is contained in m spells (i.e. in case of partially or fully overlapping spells as defined by their recorded start and end dates), I assign that day only to the spell featuring the maximum number of average contributions per day. Therefore, the day d is assigned to the spell i^* such that:
$$i^* = \underset{i \in \{1,\ldots,m\}}{\operatorname{argmax}}\ \bar{c}^d_{ijy} \tag{2.26}$$
If a day belongs to a unique spell i, then $i^* = i$.

3. After having applied this procedure to all days in a given year, I compute weeks of qualifying contributions to Social Security in that year (c_{jy}) as:
$$c_{jy} = \sum_{d=1}^{k} \bar{c}^d_{i^*jy} \tag{2.27}$$
where k takes value 365 but in leap years (when k is set to 366).

4. The number of years of contributions by 1995 (N_{1995}) is then:

$$N_{j,y_0:1995} = \frac{1}{52} \cdot \sum_{y=y_0}^{1995} c_{jy} \tag{2.28}$$

Notice that, in years that feature a single contribution spell, this procedure recovers the weeks of contributions associated to that spell; in years that features multiple but non-overlapping spells, this procedure recovers the sum of the weeks of contributions associated to those spells; if there exist two spells with the same start and end dates, the procedure recovers the weeks of contributions of the spell that features more weeks of contributions.

2.D Building measures of labour supply

To derive various measures of labour supply in any given year, I need to identify whether any single day is covered by contributions to Social Security and whether such contributions are work-related. To this end, I proceed in the following way:

1. I define d^s_{ijy} and d^e_{ijy} as the start date and the end dates of the contribution spell i of individual j in year y, as recorded in the dataset; c_{ijy} are the recorded weeks of contributions associated to that spell. I then define d^{e*}_{ijy} as the end date of the spell that would result by summing the recorded days of contributions to the recorded start date of the spell. As mentioned in Appendix Section (2.C), there are instances in which $d^{e*}_{ijy} \neq d^e_{ijy}$. To identify workers belonging to the C95 and T1-95 groups (Appendix Section 2.C) I had to rely on INPS-recorded start and end dates, because I had to follow their dispositions to manage overlapping spells. To compute measures of labour supply I instead need to identify for each day whether a contribution exists and the type of contribution.

2. I consider a day d as covered by contributions to Social Security if it belongs to any given contribution spell $[d^s_{ijy}, d^{e*}_{ijy}]$:

$$cc_{ijyd} = \mathbb{1}(d \text{ is covered by SSCs}) \text{ if } d \in [d^s_{ijy}, d^{e*}_{ijy}] \tag{2.29}$$

3. I consider a day d as covered by work-related contributions to Social Security if it belongs to any work-related contribution spell $[d^{s,W}_{ijy}, d^{e*,W}_{ijy}]$, while at the same time not being comprised also in any non-work related contribution spell $[d^{s,NW}_{ijy}, d^{e*,NW}_{ijy}]$:

$$\begin{aligned} cc^W_{ijyd} = \mathbb{1}(d \text{ is covered by work-related SSCs}) \text{ if } d \in [d^{s,W}_{ijy}, d^{e*,W}_{ijy}] \\ \text{and } d \notin [d^{s,NW}_{ijy}, d^{e*,NW}_{ijy}] \end{aligned} \tag{2.30}$$

4. The number of days covered by contributions (cc_{jy}) and the number of days covered by work-related contributions (cc^W_{jy}) in any given year are then computed as:

$$cc_{jy} = \sum_{d=1}^{k} cc_{ijyd} \tag{2.31}$$

$$cc^W_{jy} = \sum_{d=1}^{k} cc^W_{ijyd} \tag{2.32}$$

where k takes value 365 but in leap years (when k is set to 366).

2.E Additional figures and tables

Table 2.E.1: Main results

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(\text{Year} = 1990)$	0.008**	0.657	0.003	0.270	78.123
	(0.004)	(1.402)	(0.004)	(1.512)	(123.357)
$T \cdot \mathbb{1}(\text{Year} = 1991)$	0.005	-0.179	0.001	-0.325	70.178
	(0.004)	(1.306)	(0.004)	(1.415)	(117.503)
$T \cdot \mathbb{1}(\text{Year} = 1992)$	0.006*	1.095	0.004	1.539	123.312
	(0.003)	(1.193)	(0.004)	(1.306)	(107.113)
$T \cdot \mathbb{1}(\text{Year} = 1993)$	0.005*	1.006	0.003	0.402	87.689
	(0.003)	(1.024)	(0.003)	(1.144)	(92.470)
$T \cdot \mathbb{1}(\text{Year} = 1994)$	0.004*	0.155	0.003	0.310	77.909
	(0.002)	(0.739)	(0.002)	(0.840)	(70.139)
$T \cdot \mathbb{1}(\text{Year} = 1996)$	-0.000	-0.206	-0.001	0.041	8.756
	(0.002)	(0.769)	(0.003)	(0.848)	(82.004)
$T \cdot \mathbb{1}(\text{Year} = 1997)$	0.005	1.943*	0.005*	1.906*	163.922
	(0.003)	(1.038)	(0.003)	(1.122)	(116.690)
$T \cdot \mathbb{1}(\text{Year} = 1998)$	0.009***	2.787**	0.008**	2.965**	231.287*
	(0.003)	(1.169)	(0.004)	(1.251)	(125.219)
$T \cdot \mathbb{1}(\text{Year} = 1999)$	0.008**	3.098**	0.008**	3.782***	228.954*
	(0.004)	(1.245)	(0.004)	(1.318)	(135.536)
$T \cdot \mathbb{1}(\text{Year} = 2000)$	0.010***	4.390***	0.011***	4.790***	299.741**
	(0.004)	(1.307)	(0.004)	(1.384)	(142.512)
$T \cdot \mathbb{1}(\text{Year} = 2001)$	0.014***	4.873***	0.014***	5.140***	307.309**
	(0.004)	(1.364)	(0.004)	(1.448)	(151.055)
$T \cdot \mathbb{1}(\text{Year} = 2002)$	0.015***	5.509***	0.015***	5.583***	289.090*
	(0.004)	(1.411)	(0.004)	(1.503)	(155.908)
$T \cdot \mathbb{1}(\text{Year} = 2003)$	0.013***	4.951***	0.015***	5.610***	341.054**
	(0.004)	(1.458)	(0.004)	(1.553)	(162.605)
$T \cdot \mathbb{1}(\text{Year} = 2004)$	0.012***	4.183***	0.013***	4.509***	422.644**
	(0.004)	(1.494)	(0.004)	(1.587)	(166.725)
$T \cdot \mathbb{1}(\text{Year} = 2005)$	0.016***	5.341***	0.017***	5.455***	386.227**
	(0.004)	(1.525)	(0.004)	(1.622)	(172.975)
$T \cdot \mathbb{1}(\text{Year} = 2006)$	0.015***	4.588***	0.017***	5.549***	296.183
	(0.004)	(1.559)	(0.005)	(1.655)	(183.434)
Observations	543507	543507	543507	543507	538773
R-squared	0.586	0.627	0.592	0.628	0.767

Notes: The table reports estimates of the coefficients β_ks from specification (2.10), which includes individual and year fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Observations lying in the 4-week window on either side of the 18-year threshold are excluded. Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.2: Main results, including observations very close to the threshold

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(\text{Year} = 1990)$	0.009**	1.476	0.003	0.759	91.983
	(0.004)	(1.333)	(0.004)	(1.435)	(117.320)
$T \cdot \mathbb{1}(\text{Year} = 1991)$	0.007*	0.603	0.002	0.227	78.747
	(0.003)	(1.244)	(0.004)	(1.345)	(111.838)
$T \cdot \mathbb{1}(\text{Year} = 1992)$	0.007**	1.564	0.004	1.830	122.633
	(0.003)	(1.137)	(0.003)	(1.242)	(102.412)
$T \cdot \mathbb{1}(\text{Year} = 1993)$	0.006**	1.474	0.003	0.687	64.540
	(0.003)	(0.973)	(0.003)	(1.086)	(88.460)
$T \cdot \mathbb{1}(\text{Year} = 1994)$	0.005**	0.446	0.002	0.403	52.473
	(0.002)	(0.703)	(0.002)	(0.796)	(66.884)
$T \cdot \mathbb{1}(\text{Year} = 1996)$	0.000	-0.112	-0.001	0.124	6.004
	(0.002)	(0.732)	(0.003)	(0.806)	(78.594)
$T \cdot \mathbb{1}(\text{Year} = 1997)$	0.005*	2.216**	0.005	2.190**	183.652*
	(0.003)	(0.990)	(0.003)	(1.070)	(111.395)
$T \cdot \mathbb{1}(\text{Year} = 1998)$	0.009***	3.049***	0.007**	2.990**	192.668
	(0.003)	(1.114)	(0.003)	(1.191)	(119.558)
$T \cdot \mathbb{1}(\text{Year} = 1999)$	0.008**	3.464***	0.007**	3.900***	181.650
	(0.003)	(1.190)	(0.004)	(1.258)	(129.507)
$T \cdot \mathbb{1}(\text{Year} = 2000)$	0.010***	4.499***	0.011***	4.854***	273.103**
	(0.003)	(1.247)	(0.004)	(1.320)	(135.714)
$T \cdot \mathbb{1}(\text{Year} = 2001)$	0.013***	5.147***	0.014***	5.444***	338.128**
	(0.004)	(1.303)	(0.004)	(1.382)	(144.072)
$T \cdot \mathbb{1}(\text{Year} = 2002)$	0.014***	5.632***	0.014***	5.764***	319.154**
	(0.004)	(1.349)	(0.004)	(1.437)	(148.610)
$T \cdot \mathbb{1}(\text{Year} = 2003)$	0.013***	5.261***	0.014***	5.834***	371.285**
	(0.004)	(1.394)	(0.004)	(1.484)	(154.899)
$T \cdot \mathbb{1}(\text{Year} = 2004)$	0.013***	4.721***	0.013***	4.984***	436.591***
	(0.004)	(1.429)	(0.004)	(1.516)	(159.265)
$T \cdot \mathbb{1}(\text{Year} = 2005)$	0.016***	5.880***	0.017***	5.953***	391.024**
	(0.004)	(1.457)	(0.004)	(1.548)	(165.181)
$T \cdot \mathbb{1}(\text{Year} = 2006)$	0.017***	5.674***	0.018***	6.434***	343.818**
	(0.004)	(1.491)	(0.004)	(1.582)	(175.027)
Observations	597431	597431	597431	597431	592095
R-squared	0.585	0.626	0.591	0.626	0.768

Notes: The table reports estimates of the coefficients β_ks from specification (2.10), which includes individual and year fixed effects. The outcomes are: (1) a dummy that takes value 1 if at least one day in a given year is covered by SS contributions; (2) the number of days per year covered by SS contributions; (3) a dummy that takes value 1 if at least one day in a given year is covered by work-related SS contributions; (4) the number of days per year covered by work-related SS contributions; (5) yearly labour earnings. Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.3: Placebo threshold at 19 years of SS contributions

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(\text{Year} = 1990)$	0.011***	3.803***	0.012***	3.702***	244.198**
	(0.003)	(1.291)	(0.004)	(1.411)	(122.374)
$T \cdot \mathbb{1}(\text{Year} = 1991)$	0.011***	3.797***	0.011***	2.893**	77.109
	(0.003)	(1.210)	(0.003)	(1.336)	(117.489)
$T \cdot \mathbb{1}(\text{Year} = 1992)$	0.005*	1.531	0.004	0.800	-16.488
	(0.003)	(1.106)	(0.003)	(1.235)	(108.072)
$T \cdot \mathbb{1}(\text{Year} = 1993)$	-0.001	-0.029	0.001	1.322	-26.477
	(0.003)	(0.953)	(0.003)	(1.095)	(95.062)
$T \cdot \mathbb{1}(\text{Year} = 1994)$	-0.001	0.124	0.001	1.140	-11.467
	(0.002)	(0.694)	(0.002)	(0.812)	(70.569)
$T \cdot \mathbb{1}(\text{Year} = 1996)$	0.002	0.602	0.003	0.430	67.734
	(0.002)	(0.736)	(0.003)	(0.821)	(80.849)
$T \cdot \mathbb{1}(\text{Year} = 1997)$	0.003	1.089	0.006*	1.852*	95.958
	(0.003)	(0.982)	(0.003)	(1.076)	(110.962)
$T \cdot \mathbb{1}(\text{Year} = 1998)$	0.003	1.487	0.005	2.093*	85.517
	(0.003)	(1.103)	(0.003)	(1.193)	(122.031)
$T \cdot \mathbb{1}(\text{Year} = 1999)$	0.001	0.885	0.003	1.609	55.259
	(0.003)	(1.178)	(0.003)	(1.263)	(156.278)
$T \cdot \mathbb{1}(\text{Year} = 2000)$	0.002	0.536	0.004	1.201	72.494
	(0.003)	(1.241)	(0.004)	(1.326)	(137.911)
$T \cdot \mathbb{1}(\text{Year} = 2001)$	0.003	1.492	0.006	1.987	-11.504
	(0.004)	(1.303)	(0.004)	(1.391)	(147.107)
$T \cdot \mathbb{1}(\text{Year} = 2002)$	-0.001	0.666	0.001	1.003	-5.778
	(0.004)	(1.343)	(0.004)	(1.437)	(150.827)
$T \cdot \mathbb{1}(\text{Year} = 2003)$	-0.000	0.737	0.002	0.486	-106.467
	(0.004)	(1.389)	(0.004)	(1.492)	(159.489)
$T \cdot \mathbb{1}(\text{Year} = 2004)$	0.001	1.041	0.003	1.796	-82.849
	(0.004)	(1.428)	(0.004)	(1.533)	(163.763)
$T \cdot \mathbb{1}(\text{Year} = 2005)$	0.000	0.838	0.004	2.217	-55.723
	(0.004)	(1.455)	(0.004)	(1.566)	(170.758)
$T \cdot \mathbb{1}(\text{Year} = 2006)$	-0.004	0.711	-0.001	2.030	-83.943
	(0.004)	(1.489)	(0.004)	(1.600)	(181.213)
Observations	540056	540056	540056	540056	535730
R-squared	0.573	0.614	0.581	0.615	0.763

Notes: The table reports estimates of coefficients β_ks from specification (2.10), which includes individual and year fixed effects. The placebo threshold of the *Dini* reform that separates fully grandfathered from partially grandfathered individuals is set at 19 years of contributions, rather than at 18. See notes to table 2.E.1 for details on the sample and the outcomes.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.4: Placebo threshold at 20 years of SS contributions

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(\text{Year} = 1990)$	0.005	1.328	0.003	0.877	122.339
	(0.003)	(1.237)	(0.004)	(1.367)	(122.103)
$T \cdot \mathbb{1}(\text{Year} = 1991)$	0.003	1.019	-0.000	0.980	149.954
	(0.003)	(1.166)	(0.003)	(1.309)	(118.391)
$T \cdot \mathbb{1}(\text{Year} = 1992)$	0.003	0.931	0.002	0.618	91.558
	(0.003)	(1.066)	(0.003)	(1.207)	(108.533)
$T \cdot \mathbb{1}(\text{Year} = 1993)$	0.002	-0.305	0.001	-0.751	119.421
	(0.002)	(0.931)	(0.003)	(1.084)	(94.489)
$T \cdot \mathbb{1}(\text{Year} = 1994)$	0.001	-0.066	-0.001	0.151	92.158
	(0.002)	(0.688)	(0.002)	(0.804)	(70.797)
$T \cdot \mathbb{1}(\text{Year} = 1996)$	0.001	0.401	0.001	0.701	-14.383
	(0.002)	(0.714)	(0.002)	(0.806)	(80.552)
$T \cdot \mathbb{1}(\text{Year} = 1997)$	0.002	0.298	0.001	0.351	146.072
	(0.003)	(0.962)	(0.003)	(1.062)	(107.115)
$T \cdot \mathbb{1}(\text{Year} = 1998)$	-0.001	-0.769	-0.002	-0.554	137.386
	(0.003)	(1.074)	(0.003)	(1.167)	(120.465)
$T \cdot \mathbb{1}(\text{Year} = 1999)$	0.001	0.593	0.001	0.234	134.530
	(0.003)	(1.153)	(0.003)	(1.240)	(154.808)
$T \cdot \mathbb{1}(\text{Year} = 2000)$	-0.000	0.172	-0.002	-0.369	11.008
	(0.003)	(1.205)	(0.003)	(1.293)	(139.050)
$T \cdot \mathbb{1}(\text{Year} = 2001)$	0.000	-0.042	-0.001	0.229	60.946
	(0.003)	(1.258)	(0.004)	(1.351)	(146.615)
$T \cdot \mathbb{1}(\text{Year} = 2002)$	0.002	0.751	0.001	1.202	180.991
	(0.004)	(1.296)	(0.004)	(1.397)	(150.329)
$T \cdot \mathbb{1}(\text{Year} = 2003)$	0.002	0.397	-0.001	0.578	253.305
	(0.004)	(1.330)	(0.004)	(1.439)	(158.463)
$T \cdot \mathbb{1}(\text{Year} = 2004)$	0.000	0.483	0.001	1.321	208.595
	(0.004)	(1.374)	(0.004)	(1.498)	(164.790)
$T \cdot \mathbb{1}(\text{Year} = 2005)$	0.005	1.531	0.006	2.398	312.811*
	(0.004)	(1.413)	(0.004)	(1.542)	(174.760)
$T \cdot \mathbb{1}(\text{Year} = 2006)$	0.007*	2.102	0.008*	3.267**	378.568**
	(0.004)	(1.447)	(0.004)	(1.579)	(183.955)
Observations	521815	521815	521815	521815	517568
R-squared	0.563	0.602	0.569	0.603	0.764

Notes: The table reports estimates of coefficients β_ks from specification (2.10), which includes individual and year fixed effects. The placebo threshold of the *Dini* reform that separates fully grandfathered from partially grandfathered individuals is set at 20 years of contributions, rather than at 18. See notes to table 2.E.1 for details on the sample and the outcomes.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.5: Main results, heterogeneity by gender

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(Year = 1990)$	0.020***	2.166	0.013*	0.776	11.860
	(0.007)	(2.672)	(0.008)	(2.883)	(179.278)
$T \cdot \mathbb{1}(Year = 1991)$	0.012*	1.150	0.005	1.395	-58.672
	(0.007)	(2.477)	(0.007)	(2.675)	(171.945)
$T \cdot \mathbb{1}(Year = 1992)$	0.014**	1.731	0.009	1.781	-67.726
	(0.006)	(2.208)	(0.007)	(2.403)	(151.675)
$T \cdot \mathbb{1}(Year = 1993)$	0.004	-0.751	0.002	-2.259	-170.003
	(0.006)	(1.850)	(0.006)	(2.030)	(132.859)
$T \cdot \mathbb{1}(Year = 1994)$	0.007*	-1.074	0.004	-1.561	-139.231
	(0.004)	(1.297)	(0.004)	(1.466)	(96.783)
$T \cdot \mathbb{1}(Year = 1996)$	0.008*	2.313*	0.010**	2.927**	101.169
	(0.004)	(1.291)	(0.005)	(1.433)	(102.282)
$T \cdot \mathbb{1}(Year = 1997)$	0.016***	5.444***	0.017***	4.846**	122.468
	(0.005)	(1.764)	(0.006)	(1.929)	(141.012)
$T \cdot \mathbb{1}(Year = 1998)$	0.019***	5.160**	0.019***	5.340**	47.450
	(0.006)	(2.029)	(0.006)	(2.185)	(160.310)
$T \cdot \mathbb{1}(Year = 1999)$	0.021***	6.540***	0.022***	6.859***	120.434
	(0.007)	(2.221)	(0.007)	(2.337)	(173.144)
$T \cdot \mathbb{1}(Year = 2000)$	0.021***	8.191***	0.023***	8.262***	248.617
	(0.007)	(2.362)	(0.007)	(2.477)	(184.497)
$T \cdot \mathbb{1}(Year = 2001)$	0.028***	9.273***	0.030***	9.492***	303.967
	(0.007)	(2.493)	(0.007)	(2.608)	(189.880)
$T \cdot \mathbb{1}(Year = 2002)$	0.034***	11.387***	0.036***	11.315***	393.919*
	(0.008)	(2.592)	(0.008)	(2.721)	(203.079)
$T \cdot \mathbb{1}(Year = 2003)$	0.034***	11.819***	0.037***	12.460***	517.846**
	(0.008)	(2.697)	(0.008)	(2.826)	(206.747)
$T \cdot \mathbb{1}(Year = 2004)$	0.029***	9.183***	0.033***	9.728***	418.605**
	(0.008)	(2.772)	(0.008)	(2.897)	(212.129)
$T \cdot \mathbb{1}(Year = 2005)$	0.034***	10.695***	0.040***	11.121***	375.586*
	(0.008)	(2.834)	(0.008)	(2.963)	(221.165)
$T \cdot \mathbb{1}(Year = 2006)$	0.031***	9.477***	0.037***	10.925***	162.728
	(0.008)	(2.902)	(0.009)	(3.039)	(236.154)
$T \cdot \mathbb{1}(Year = 1990) \cdot Male$	-0.021**	-2.752	-0.017*	-0.927	77.506
	(0.008)	(3.100)	(0.009)	(3.347)	(243.546)
$T \cdot \mathbb{1}(Year = 1991) \cdot Male$	-0.012	-2.399	-0.007	-2.796	191.983
	(0.008)	(2.881)	(0.008)	(3.119)	(232.690)
$T \cdot \mathbb{1}(Year = 1992) \cdot Male$	-0.013*	-1.226	-0.010	-0.411	298.307
	(0.007)	(2.603)	(0.008)	(2.844)	(209.727)
$T \cdot \mathbb{1}(Year = 1993) \cdot Male$	0.002	2.627	0.002	4.131*	402.981**
	(0.006)	(2.209)	(0.007)	(2.448)	(181.944)
$T \cdot \mathbb{1}(Year = 1994) \cdot Male$	-0.005	1.890	-0.002	2.918	341.803**
	(0.005)	(1.574)	(0.005)	(1.785)	(136.075)
$T \cdot \mathbb{1}(Year = 1996) \cdot Male$	-0.013**	-3.916**	-0.017***	-4.533**	-135.995
	(0.005)	(1.607)	(0.006)	(1.778)	(154.108)
$T \cdot \mathbb{1}(Year = 1997) \cdot Male$	-0.017***	-5.397**	-0.018**	-4.524*	99.061
	(0.006)	(2.179)	(0.007)	(2.368)	(217.254)
$T \cdot \mathbb{1}(Year = 1998) \cdot Male$	-0.015**	-3.587	-0.018**	-3.617	328.193
	(0.007)	(2.477)	(0.008)	(2.659)	(236.898)
$T \cdot \mathbb{1}(Year = 1999) \cdot Male$	-0.020***	-5.209*	-0.022***	-4.676*	218.331
	(0.008)	(2.671)	(0.008)	(2.820)	(256.113)
$T \cdot \mathbb{1}(Year = 2000) \cdot Male$	-0.016**	-5.735**	-0.018**	-5.269*	131.889
	(0.008)	(2.819)	(0.008)	(2.973)	(270.345)
$T \cdot \mathbb{1}(Year = 2001) \cdot Male$	-0.021**	-6.648**	-0.024***	-6.640**	59.301
	(0.008)	(2.957)	(0.009)	(3.119)	(284.075)
$T \cdot \mathbb{1}(Year = 2002) \cdot Male$	-0.029***	-8.932***	-0.033***	-8.778***	-113.896
	(0.009)	(3.066)	(0.009)	(3.245)	(296.399)
$T \cdot \mathbb{1}(Year = 2003) \cdot Male$	-0.032***	-10.489***	-0.034***	-10.514***	-220.321
	(0.009)	(3.179)	(0.009)	(3.361)	(306.688)
$T \cdot \mathbb{1}(Year = 2004) \cdot Male$	-0.026***	-7.495**	-0.032***	-7.898**	74.400
	(0.009)	(3.261)	(0.010)	(3.439)	(314.416)
$T \cdot \mathbb{1}(Year = 2005) \cdot Male$	-0.026***	-8.059**	-0.034***	-8.597**	86.959
	(0.009)	(3.332)	(0.010)	(3.516)	(326.788)
$T \cdot \mathbb{1}(Year = 2006) \cdot Male$	-0.023**	-7.276**	-0.032***	-8.080**	296.485
	(0.010)	(3.408)	(0.010)	(3.596)	(347.004)
Observations	543507	543507	543507	543507	538773
R-squared	0.589	0.629	0.594	0.628	0.768

Notes: The table shows coefficients β_ks and β_k^ds from specification (2.12), which includes individual and year-gender fixed effects. See notes to Figure 2.9 for further details on the outcomes and the sample.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.6: Main results, heterogeneity by age in 1995

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(Year = 1990)$	0.016***	5.380***	0.010**	2.741	-199.898
	(0.004)	(1.799)	(0.005)	(2.139)	(181.725)
$T \cdot \mathbb{1}(Year = 1991)$	0.013***	3.631**	0.007	2.433	-128.379
	(0.004)	(1.721)	(0.005)	(2.067)	(176.562)
$T \cdot \mathbb{1}(Year = 1992)$	0.014***	4.570***	0.009*	3.250*	2.180
	(0.004)	(1.631)	(0.005)	(1.970)	(162.716)
$T \cdot \mathbb{1}(Year = 1993)$	0.008**	2.271	0.002	0.962	79.693
	(0.004)	(1.510)	(0.004)	(1.840)	(144.838)
$T \cdot \mathbb{1}(Year = 1994)$	0.003	0.662	0.002	0.555	-73.162
	(0.003)	(1.098)	(0.004)	(1.360)	(111.362)
$T \cdot \mathbb{1}(Year = 1996)$	-0.000	-0.167	-0.001	-0.513	108.675
	(0.003)	(1.118)	(0.004)	(1.343)	(124.778)
$T \cdot \mathbb{1}(Year = 1997)$	0.003	1.430	0.001	-0.302	333.717*
	(0.004)	(1.507)	(0.005)	(1.750)	(175.198)
$T \cdot \mathbb{1}(Year = 1998)$	0.008*	1.960	0.002	0.277	320.566*
	(0.005)	(1.708)	(0.005)	(1.917)	(188.871)
$T \cdot \mathbb{1}(Year = 1999)$	0.008	2.686	0.004	1.747	370.404*
	(0.005)	(1.801)	(0.005)	(1.995)	(203.672)
$T \cdot \mathbb{1}(Year = 2000)$	0.010**	3.695*	0.008	2.041	511.119**
	(0.005)	(1.897)	(0.005)	(2.091)	(216.334)
$T \cdot \mathbb{1}(Year = 2001)$	0.011**	3.989**	0.007	2.665	534.378**
	(0.005)	(1.959)	(0.006)	(2.176)	(229.151)
$T \cdot \mathbb{1}(Year = 2002)$	0.014***	5.752***	0.012*	4.813**	806.136***
	(0.005)	(2.045)	(0.006)	(2.262)	(234.788)
$T \cdot \mathbb{1}(Year = 2003)$	0.015***	5.885***	0.015**	5.888**	1,019.041***
	(0.006)	(2.104)	(0.006)	(2.321)	(246.207)
$T \cdot \mathbb{1}(Year = 2004)$	0.012**	4.344**	0.011*	4.330*	858.189***
	(0.006)	(2.130)	(0.006)	(2.348)	(258.453)
$T \cdot \mathbb{1}(Year = 2005)$	0.007	3.801*	0.006	3.568	767.319***
	(0.006)	(2.180)	(0.006)	(2.396)	(271.706)
$T \cdot \mathbb{1}(Year = 2006)$	0.010*	2.866	0.008	2.906	856.967***
	(0.006)	(2.217)	(0.006)	(2.429)	(285.022)
$T \cdot \mathbb{1}(Year = 1990) \cdot$ age 38-45	-0.010	-6.408**	-0.009	-3.019	485.314**
	(0.007)	(2.629)	(0.008)	(2.937)	(243.468)
$T \cdot \mathbb{1}(Year = 1991) \cdot$ age 38-45	-0.010	-5.226**	-0.007	-3.696	349.909
	(0.006)	(2.473)	(0.007)	(2.785)	(233.941)
$T \cdot \mathbb{1}(Year = 1992) \cdot$ age 38-45	-0.010*	-4.925**	-0.007	-2.246	222.137
	(0.006)	(2.291)	(0.007)	(2.606)	(214.314)
$T \cdot \mathbb{1}(Year = 1993) \cdot$ age 38-45	-0.004	-1.810	0.001	-0.697	33.914
	(0.005)	(2.024)	(0.006)	(2.344)	(187.358)
$T \cdot \mathbb{1}(Year = 1994) \cdot$ age 38-45	0.002	-0.770	0.001	-0.368	234.263
	(0.004)	(1.465)	(0.005)	(1.727)	(142.917)
$T \cdot \mathbb{1}(Year = 1996) \cdot$ age 38-45	0.001	0.077	0.000	0.982	-141.839
	(0.005)	(1.511)	(0.005)	(1.727)	(164.124)
$T \cdot \mathbb{1}(Year = 1997) \cdot$ age 38-45	0.004	1.046	0.007	3.642	-244.242
	(0.006)	(2.039)	(0.006)	(2.270)	(232.249)
$T \cdot \mathbb{1}(Year = 1998) \cdot$ age 38-45	0.001	1.592	0.009	4.416*	-120.287
	(0.006)	(2.302)	(0.007)	(2.512)	(249.617)
$T \cdot \mathbb{1}(Year = 1999) \cdot$ age 38-45	0.001	0.924	0.007	3.394	-203.209
	(0.007)	(2.443)	(0.007)	(2.634)	(270.005)
$T \cdot \mathbb{1}(Year = 2000) \cdot$ age 38-45	0.001	1.366	0.006	4.484	-309.423
	(0.007)	(2.567)	(0.008)	(2.763)	(285.093)
$T \cdot \mathbb{1}(Year = 2001) \cdot$ age 38-45	0.006	1.682	0.012	4.138	-330.823
	(0.007)	(2.668)	(0.008)	(2.884)	(301.930)
$T \cdot \mathbb{1}(Year = 2002) \cdot$ age 38-45	0.002	0.003	0.006	1.554	-774.455**
	(0.008)	(2.770)	(0.008)	(2.996)	(310.617)
$T \cdot \mathbb{1}(Year = 2003) \cdot$ age 38-45	-0.003	-1.078	0.002	-0.063	-1,021.691***
	(0.008)	(2.858)	(0.008)	(3.087)	(324.627)
$T \cdot \mathbb{1}(Year = 2004) \cdot$ age 38-45	0.001	0.026	0.004	0.581	-665.242**
	(0.008)	(2.914)	(0.008)	(3.141)	(336.221)
$T \cdot \mathbb{1}(Year = 2005) \cdot$ age 38-45	0.015*	2.595	0.018**	3.144	-585.712*
	(0.008)	(2.978)	(0.009)	(3.207)	(350.808)
$T \cdot \mathbb{1}(Year = 2006) \cdot$ age 38-45	0.008	2.750	0.014	4.185	-878.637**
	(0.008)	(3.038)	(0.009)	(3.265)	(370.059)
Observations	543507	543507	543507	543507	538773
R-squared	0.586	0.628	0.593	0.628	0.767

Notes: The table shows coefficients β_ks and β_k^ds from specification (2.12), which includes individual and year-age group fixed effects. See notes to Figure 2.10 for further details on the outcomes and the sample.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.7: Main results, heterogeneity by (a proxy of) education

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(Year = 1990)$	0.012***	2.189	0.006	1.004	80.327
	(0.004)	(1.640)	(0.005)	(1.759)	(141.223)
$T \cdot \mathbb{1}(Year = 1991)$	0.008**	0.450	0.004	-0.031	36.063
	(0.004)	(1.520)	(0.004)	(1.640)	(134.429)
$T \cdot \mathbb{1}(Year = 1992)$	0.009**	1.911	0.006	2.073	95.451
	(0.004)	(1.384)	(0.004)	(1.513)	(122.052)
$T \cdot \mathbb{1}(Year = 1993)$	0.006*	1.044	0.001	-0.051	24.718
	(0.003)	(1.184)	(0.004)	(1.318)	(104.473)
$T \cdot \mathbb{1}(Year = 1994)$	0.003	-0.330	0.002	-0.334	6.029
	(0.003)	(0.840)	(0.003)	(0.957)	(78.646)
$T \cdot \mathbb{1}(Year = 1996)$	0.000	0.051	0.001	0.573	70.315
	(0.003)	(0.867)	(0.003)	(0.960)	(92.537)
$T \cdot \mathbb{1}(Year = 1997)$	0.007**	2.618**	0.007*	2.431*	201.750
	(0.003)	(1.177)	(0.004)	(1.273)	(131.194)
$T \cdot \mathbb{1}(Year = 1998)$	0.009**	3.290**	0.009**	3.394**	271.593*
	(0.004)	(1.325)	(0.004)	(1.415)	(143.642)
$T \cdot \mathbb{1}(Year = 1999)$	0.010**	3.818***	0.010**	4.022***	209.395
	(0.004)	(1.408)	(0.004)	(1.487)	(155.230)
$T \cdot \mathbb{1}(Year = 2000)$	0.010**	4.211***	0.011**	3.915**	255.724
	(0.004)	(1.481)	(0.004)	(1.568)	(164.086)
$T \cdot \mathbb{1}(Year = 2001)$	0.012***	4.396***	0.011**	4.312***	241.307
	(0.004)	(1.542)	(0.005)	(1.632)	(174.012)
$T \cdot \mathbb{1}(Year = 2002)$	0.013***	4.885***	0.012***	4.461***	215.296
	(0.004)	(1.598)	(0.005)	(1.694)	(178.698)
$T \cdot \mathbb{1}(Year = 2003)$	0.012***	4.632***	0.014***	4.716***	317.746*
	(0.005)	(1.648)	(0.005)	(1.746)	(185.562)
$T \cdot \mathbb{1}(Year = 2004)$	0.009**	3.603**	0.010**	3.505**	317.947*
	(0.005)	(1.683)	(0.005)	(1.779)	(190.656)
$T \cdot \mathbb{1}(Year = 2005)$	0.013***	4.500***	0.014***	4.413**	251.178
	(0.005)	(1.715)	(0.005)	(1.819)	(197.911)
$T \cdot \mathbb{1}(Year = 2006)$	0.012**	3.732**	0.013**	4.351**	230.509
	(0.005)	(1.749)	(0.005)	(1.851)	(210.088)
$T \cdot \mathbb{1}(Year = 1990) \cdot$ high edu.	-0.015*	-5.087	-0.011	-3.190	-4.009
	(0.008)	(3.221)	(0.009)	(3.528)	(305.959)
$T \cdot \mathbb{1}(Year = 1991) \cdot$ high edu.	-0.007	-1.810	-0.009	-1.354	83.467
	(0.008)	(3.058)	(0.009)	(3.334)	(290.919)
$T \cdot \mathbb{1}(Year = 1992) \cdot$ high edu.	-0.009	-2.904	-0.012	-2.867	12.845
	(0.007)	(2.794)	(0.008)	(3.075)	(266.799)
$T \cdot \mathbb{1}(Year = 1993) \cdot$ high edu.	0.000	0.015	0.007	1.335	171.932
	(0.007)	(2.430)	(0.008)	(2.729)	(235.155)
$T \cdot \mathbb{1}(Year = 1994) \cdot$ high edu.	0.004	2.187	0.002	2.514	269.208
	(0.005)	(1.837)	(0.006)	(2.076)	(182.807)
$T \cdot \mathbb{1}(Year = 1996) \cdot$ high edu.	0.001	-0.135	-0.008	-1.117	-145.361
	(0.006)	(1.937)	(0.007)	(2.112)	(213.057)
$T \cdot \mathbb{1}(Year = 1997) \cdot$ high edu.	-0.004	-0.582	-0.001	-0.105	31.247
	(0.007)	(2.596)	(0.008)	(2.807)	(303.474)
$T \cdot \mathbb{1}(Year = 1998) \cdot$ high edu.	0.004	0.973	0.002	1.037	95.859
	(0.008)	(2.929)	(0.009)	(3.154)	(308.447)
$T \cdot \mathbb{1}(Year = 1999) \cdot$ high edu.	-0.003	-0.452	-0.002	1.128	293.990
	(0.009)	(3.133)	(0.009)	(3.344)	(333.769)
$T \cdot \mathbb{1}(Year = 2000) \cdot$ high edu.	0.005	2.933	0.008	5.845*	445.105
	(0.009)	(3.293)	(0.010)	(3.505)	(347.175)
$T \cdot \mathbb{1}(Year = 2001) \cdot$ high edu.	0.013	5.194	0.019*	6.469*	608.662*
	(0.009)	(3.469)	(0.010)	(3.700)	(369.637)
$T \cdot \mathbb{1}(Year = 2002) \cdot$ high edu.	0.011	4.903	0.015	7.467*	652.080*
	(0.010)	(3.548)	(0.010)	(3.825)	(386.871)
$T \cdot \mathbb{1}(Year = 2003) \cdot$ high edu.	0.009	4.168	0.014	7.194*	488.744
	(0.010)	(3.687)	(0.011)	(3.978)	(409.704)
$T \cdot \mathbb{1}(Year = 2004) \cdot$ high edu.	0.020*	5.878	0.022**	7.821*	899.426**
	(0.010)	(3.813)	(0.011)	(4.090)	(416.848)
$T \cdot \mathbb{1}(Year = 2005) \cdot$ high edu.	0.020*	7.253*	0.022*	8.334**	1,069.746**
	(0.011)	(3.910)	(0.011)	(4.177)	(431.152)
$T \cdot \mathbb{1}(Year = 2006) \cdot$ high edu.	0.024**	7.277*	0.025**	8.471**	771.465*
	(0.011)	(4.009)	(0.012)	(4.279)	(455.667)
Observations	532746	532746	532746	532746	528251
R-squared	0.589	0.630	0.595	0.629	0.767

Notes: The table shows coefficients β_ks and β_k^ds from specification (2.12), which includes individual and year-education group fixed effects. See notes to Figure 2.11 for further details on the outcomes and the sample.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.8: Main results, heterogeneity by pre-1996 labour earnings

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(Year = 1990)$	0.006	-1.580	0.000	-1.962	-157.512
	(0.005)	(1.964)	(0.006)	(2.105)	(127.159)
$T \cdot \mathbb{1}(Year = 1991)$	0.004	-2.138	-0.001	-1.949	-225.635*
	(0.005)	(1.832)	(0.005)	(1.971)	(125.141)
$T \cdot \mathbb{1}(Year = 1992)$	0.005	0.037	0.001	0.542	-128.091
	(0.005)	(1.679)	(0.005)	(1.826)	(115.038)
$T \cdot \mathbb{1}(Year = 1993)$	0.007*	0.770	0.003	0.032	-115.284
	(0.004)	(1.439)	(0.004)	(1.594)	(100.868)
$T \cdot \mathbb{1}(Year = 1994)$	0.006*	0.308	0.004	0.705	-47.544
	(0.003)	(1.040)	(0.004)	(1.174)	(75.771)
$T \cdot \mathbb{1}(Year = 1996)$	0.002	0.313	0.001	0.278	-2.062
	(0.003)	(1.079)	(0.004)	(1.182)	(87.938)
$T \cdot \mathbb{1}(Year = 1997)$	0.008*	2.797*	0.007	2.388	174.998
	(0.004)	(1.441)	(0.004)	(1.542)	(126.294)
$T \cdot \mathbb{1}(Year = 1998)$	0.013***	3.787**	0.011**	3.737**	266.913*
	(0.005)	(1.616)	(0.005)	(1.716)	(137.952)
$T \cdot \mathbb{1}(Year = 1999)$	0.010**	3.836**	0.008*	4.116**	261.702*
	(0.005)	(1.727)	(0.005)	(1.816)	(150.382)
$T \cdot \mathbb{1}(Year = 2000)$	0.015***	5.663***	0.015***	5.921***	322.433**
	(0.005)	(1.806)	(0.005)	(1.906)	(155.144)
$T \cdot \mathbb{1}(Year = 2001)$	0.017***	5.371***	0.017***	5.602***	294.372*
	(0.005)	(1.882)	(0.006)	(1.989)	(163.062)
$T \cdot \mathbb{1}(Year = 2002)$	0.019***	6.012***	0.017***	5.970***	267.733
	(0.005)	(1.946)	(0.006)	(2.065)	(164.999)
$T \cdot \mathbb{1}(Year = 2003)$	0.016***	5.362***	0.017***	6.058***	275.670
	(0.006)	(2.010)	(0.006)	(2.133)	(167.762)
$T \cdot \mathbb{1}(Year = 2004)$	0.016***	4.775**	0.015**	4.970**	341.267**
	(0.006)	(2.060)	(0.006)	(2.178)	(170.601)
$T \cdot \mathbb{1}(Year = 2005)$	0.020***	6.930***	0.020***	6.856***	390.748**
	(0.006)	(2.102)	(0.006)	(2.224)	(173.843)
$T \cdot \mathbb{1}(Year = 2006)$	0.022***	6.204***	0.022***	7.146***	392.627**
	(0.006)	(2.143)	(0.006)	(2.263)	(184.216)
$T \cdot \mathbb{1}(Year = 1990) \cdot$ top terc.	-0.005	1.798	0.000	2.070	-93.114
	(0.006)	(2.332)	(0.007)	(2.579)	(277.526)
$T \cdot \mathbb{1}(Year = 1991) \cdot$ top terc.	-0.002	2.203	0.002	1.699	334.245
	(0.006)	(2.189)	(0.006)	(2.437)	(269.461)
$T \cdot \mathbb{1}(Year = 1992) \cdot$ top terc.	-0.004	0.132	0.001	0.385	305.240
	(0.005)	(2.003)	(0.006)	(2.258)	(247.402)
$T \cdot \mathbb{1}(Year = 1993) \cdot$ top terc.	-0.008*	-1.362	-0.005	-0.570	260.916
	(0.005)	(1.742)	(0.005)	(2.027)	(212.725)
$T \cdot \mathbb{1}(Year = 1994) \cdot$ top terc.	-0.006*	-0.982	-0.005	-1.478	161.851
	(0.004)	(1.253)	(0.004)	(1.485)	(163.936)
$T \cdot \mathbb{1}(Year = 1996) \cdot$ top terc.	-0.003	-1.218	-0.003	-0.942	-50.898
	(0.004)	(1.335)	(0.005)	(1.535)	(198.982)
$T \cdot \mathbb{1}(Year = 1997) \cdot$ top terc.	-0.006	-3.058*	-0.005	-2.839	-55.285
	(0.005)	(1.811)	(0.006)	(2.034)	(283.354)
$T \cdot \mathbb{1}(Year = 1998) \cdot$ top terc.	-0.010*	-3.976*	-0.011*	-4.421*	-134.548
	(0.006)	(2.063)	(0.006)	(2.276)	(300.056)
$T \cdot \mathbb{1}(Year = 1999) \cdot$ top terc.	-0.008	-3.571	-0.004	-3.148	-104.182
	(0.006)	(2.190)	(0.006)	(2.376)	(323.019)
$T \cdot \mathbb{1}(Year = 2000) \cdot$ top terc.	-0.014**	-5.793**	-0.014**	-6.247**	-76.047
	(0.006)	(2.316)	(0.007)	(2.497)	(345.968)
$T \cdot \mathbb{1}(Year = 2001) \cdot$ top terc.	-0.012*	-3.723	-0.013*	-4.717*	-16.107
	(0.007)	(2.425)	(0.007)	(2.618)	(368.306)
$T \cdot \mathbb{1}(Year = 2002) \cdot$ top terc.	-0.013*	-4.099	-0.014*	-4.682*	-6.048
	(0.007)	(2.505)	(0.007)	(2.712)	(385.914)
$T \cdot \mathbb{1}(Year = 2003) \cdot$ top terc.	-0.010	-3.615	-0.013*	-4.446	195.438
	(0.007)	(2.602)	(0.007)	(2.809)	(407.955)
$T \cdot \mathbb{1}(Year = 2004) \cdot$ top terc.	-0.013*	-3.626	-0.012	-4.049	291.000
	(0.007)	(2.666)	(0.008)	(2.876)	(420.332)
$T \cdot \mathbb{1}(Year = 2005) \cdot$ top terc.	-0.014*	-6.765**	-0.015*	-6.943**	43.182
	(0.007)	(2.727)	(0.008)	(2.947)	(441.727)
$T \cdot \mathbb{1}(Year = 2006) \cdot$ top terc.	-0.021***	-6.681**	-0.022***	-7.726**	-188.526
	(0.008)	(2.803)	(0.008)	(3.026)	(468.856)
Observations	531845	531845	531845	531845	530008
R-squared	0.583	0.620	0.585	0.615	0.768

Notes: The table shows coefficients β_ks and β_k^ds from specification (2.12), which includes individual and year-income group fixed effects. See notes to Figure 2.12 for further details on the outcomes and the sample.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.9: Main results, heterogeneity by pre-1996 occupation

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
$T \cdot \mathbb{1}(\text{Year} = 1990)$	0.012***	1.883	0.007	1.426	143.411
	(0.004)	(1.605)	(0.005)	(1.780)	(141.388)
$T \cdot \mathbb{1}(\text{Year} = 1991)$	0.010**	1.462	0.006	1.552	150.537
	(0.004)	(1.498)	(0.004)	(1.674)	(130.110)
$T \cdot \mathbb{1}(\text{Year} = 1992)$	0.011***	2.266*	0.009**	3.118**	228.439**
	(0.004)	(1.374)	(0.004)	(1.553)	(115.797)
$T \cdot \mathbb{1}(\text{Year} = 1993)$	0.008**	2.159*	0.006	1.283	148.027
	(0.003)	(1.202)	(0.004)	(1.390)	(99.974)
$T \cdot \mathbb{1}(\text{Year} = 1994)$	0.008***	0.406	0.006*	0.365	104.436
	(0.003)	(0.877)	(0.003)	(1.037)	(73.880)
$T \cdot \mathbb{1}(\text{Year} = 1996)$	-0.000	-0.219	-0.001	0.190	-40.539
	(0.003)	(0.898)	(0.003)	(1.032)	(83.236)
$T \cdot \mathbb{1}(\text{Year} = 1997)$	0.003	1.264	0.005	1.483	34.645
	(0.003)	(1.202)	(0.004)	(1.340)	(116.204)
$T \cdot \mathbb{1}(\text{Year} = 1998)$	0.009**	2.526*	0.009**	3.149**	158.875
	(0.004)	(1.358)	(0.004)	(1.492)	(130.420)
$T \cdot \mathbb{1}(\text{Year} = 1999)$	0.009**	2.980**	0.010**	4.199***	124.575
	(0.004)	(1.446)	(0.004)	(1.566)	(142.684)
$T \cdot \mathbb{1}(\text{Year} = 2000)$	0.009**	4.080***	0.012**	5.058***	253.114*
	(0.004)	(1.509)	(0.005)	(1.636)	(152.276)
$T \cdot \mathbb{1}(\text{Year} = 2001)$	0.012***	4.956***	0.013***	5.572***	206.982
	(0.004)	(1.569)	(0.005)	(1.706)	(162.119)
$T \cdot \mathbb{1}(\text{Year} = 2002)$	0.015***	5.579***	0.015***	6.049***	138.601
	(0.005)	(1.615)	(0.005)	(1.765)	(171.064)
$T \cdot \mathbb{1}(\text{Year} = 2003)$	0.015***	5.593***	0.018***	6.642***	176.240
	(0.005)	(1.662)	(0.005)	(1.817)	(181.895)
$T \cdot \mathbb{1}(\text{Year} = 2004)$	0.014***	5.152***	0.016***	5.794***	249.418
	(0.005)	(1.704)	(0.005)	(1.856)	(187.884)
$T \cdot \mathbb{1}(\text{Year} = 2005)$	0.017***	6.082***	0.020***	6.500***	190.216
	(0.005)	(1.739)	(0.005)	(1.894)	(196.385)
$T \cdot \mathbb{1}(\text{Year} = 2006)$	0.016***	5.385***	0.020***	7.089***	84.518
	(0.005)	(1.781)	(0.005)	(1.934)	(208.589)
$T \cdot \mathbb{1}(\text{Year} = 1990) \cdot$ self-emp	-0.015*	-4.120	-0.013	-3.840	-202.824
	(0.009)	(3.225)	(0.009)	(3.365)	(283.372)
$T \cdot \mathbb{1}(\text{Year} = 1991) \cdot$ self-emp.	-0.014*	-5.550*	-0.016*	-6.328**	-257.964
	(0.008)	(3.000)	(0.008)	(3.132)	(280.183)
$T \cdot \mathbb{1}(\text{Year} = 1992) \cdot$ self-emp.	-0.016**	-3.971	-0.018**	-5.332*	-352.322
	(0.007)	(2.726)	(0.008)	(2.871)	(261.360)
$T \cdot \mathbb{1}(\text{Year} = 1993) \cdot$ self-emp.	-0.010	-3.880*	-0.011	-3.004	-225.726
	(0.006)	(2.291)	(0.007)	(2.437)	(224.798)
$T \cdot \mathbb{1}(\text{Year} = 1994) \cdot$ self-emp.	-0.013***	-0.843	-0.011**	-0.202	-103.291
	(0.005)	(1.629)	(0.005)	(1.747)	(174.439)
$T \cdot \mathbb{1}(\text{Year} = 1996) \cdot$ self-emp.	0.000	0.043	-0.001	-0.501	160.270
	(0.006)	(1.729)	(0.006)	(1.807)	(210.448)
$T \cdot \mathbb{1}(\text{Year} = 1997) \cdot$ self-emp.	0.007	2.302	0.003	1.469	425.606
	(0.007)	(2.356)	(0.007)	(2.449)	(302.686)
$T \cdot \mathbb{1}(\text{Year} = 1998) \cdot$ self-emp.	-0.000	0.927	-0.004	-0.527	233.806
	(0.007)	(2.644)	(0.008)	(2.735)	(314.743)
$T \cdot \mathbb{1}(\text{Year} = 1999) \cdot$ self-emp.	-0.003	0.463	-0.009	-1.277	345.309
	(0.008)	(2.819)	(0.008)	(2.898)	(337.995)
$T \cdot \mathbb{1}(\text{Year} = 2000) \cdot$ self-emp.	0.006	1.136	-0.001	-0.730	154.410
	(0.008)	(2.975)	(0.009)	(3.058)	(351.188)
$T \cdot \mathbb{1}(\text{Year} = 2001) \cdot$ self-emp.	0.008	-0.168	0.004	-1.246	338.238
	(0.009)	(3.120)	(0.009)	(3.208)	(370.719)
$T \cdot \mathbb{1}(\text{Year} = 2002) \cdot$ self-emp.	0.001	-0.082	-0.002	-1.301	517.773
	(0.009)	(3.241)	(0.009)	(3.339)	(374.859)
$T \cdot \mathbb{1}(\text{Year} = 2003) \cdot$ self-emp.	-0.005	-1.971	-0.010	-3.176	572.885
	(0.009)	(3.365)	(0.010)	(3.466)	(383.369)
$T \cdot \mathbb{1}(\text{Year} = 2004) \cdot$ self-emp.	-0.005	-3.051	-0.010	-4.016	602.984
	(0.009)	(3.444)	(0.010)	(3.540)	(389.939)
$T \cdot \mathbb{1}(\text{Year} = 2005) \cdot$ self-emp.	-0.003	-2.276	-0.010	-3.202	679.482*
	(0.010)	(3.515)	(0.010)	(3.623)	(401.362)
$T \cdot \mathbb{1}(\text{Year} = 2006) \cdot$ self-emp.	-0.004	-2.457	-0.009	-4.845	726.363*
	(0.010)	(3.585)	(0.010)	(3.694)	(425.170)
Observations	543507	543507	543507	543507	538773
R-squared	0.586	0.628	0.593	0.629	0.767

Notes: The table shows coefficients β_ks and β_k^ds from specification (2.12), which includes individual and year-occupation fixed effects. See notes to Figure 2.13 for further details on the outcomes and the sample.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 2.E.10: Main results, heterogeneity along all dimensions but age

	SSC (1=yes) (1)	SSC (days) (2)	work SSC (1=yes) (3)	work SSC (days) (4)	labour earnings (euros) (5)
T · Post	0.009	6.356***	0.016**	7.185***	218.152
	(0.006)	(2.358)	(0.007)	(2.443)	(194.367)
T · Post · Male	-0.009	-4.208*	-0.013*	-4.475*	6.055
	(0.006)	(2.430)	(0.007)	(2.549)	(243.534)
T · Post · High edu.	0.014**	4.952*	0.016**	5.831**	304.339
	(0.007)	(2.613)	(0.007)	(2.766)	(305.455)
T · Post · Self emp.	0.009	1.487	0.005	0.197	538.806*
	(0.007)	(2.548)	(0.007)	(2.646)	(301.423)
T · Post · Top terc.	-0.002	-3.051	-0.006	-3.746*	-58.433
	(0.005)	(2.009)	(0.006)	(2.164)	(324.295)
Observations	521611	521611	521611	521611	519824
R-squared	0.588	0.625	0.589	0.619	0.770

Notes: The table reports estimates based on an enriched version of specification (2.13), where the term $T_i \cdot Post_t$ and year fixed effects are also interacted with: a dummy that takes value 1 if the individual is male; a dummy that takes value 1 if the individual is high-educated (as measured by a proxy); a dummy that takes value 1 if the individual belongs to the top tercile of the distribution of pre-1996 labour earnings; a dummy that takes value 1 if the individual worked as a self-employed in the last occupation before 1996. Standard errors are reported in parenthesis and are clustered at the individual level.
*, **, *** denote significance at the 10, 5 and 1 per cent level, respectively.

Chapter 3

It's time to learn: school institutions and returns to instruction time

3.1 Introduction

Given the important role played by schools in the formation of human capital, scholars and policymakers have long been interested in understanding what makes a school more effective. A rich literature examines how school inputs (such as class size or instruction time) affect students' achievement. A more recent literature highlights the role of school institutions and governance (Woessmann, 2016). Indeed, a growing line of research explores the effects of granting more autonomy to schools on learning outcomes (see Section 3.2). In this paper we study if and why the effects of expanding a specific school input - instruction time - vary across public and charter schools. This paper therefore lies at the intersection of the two literatures, as we address the question of whether the way in which schools are managed and organized affects how they use additional resources.

Studying how school institutions affect returns to additional instruction time is interesting because many countries are considering devoting, or have already allocated, substantial funds to increasing the amount of time that pupils spend at school.[1] In addition, since time is an inherently limited resource, extended school schedules reduce the amount of time that students can dedicate to other activities. Therefore, the effect on achievement could depend not only on the absolute quality of time use at school, but also on its relative quality with respect to the learning opportunities available outside of school. Hence, we also examine whether the effects of increased instruction time are heterogeneous across pupils from different socio-economic backgrounds.

[1] For instance, since 2003 Germany has begun phasing in all-day schooling and the percentage of pupils attending all-day primary schools has increased from 7.9 per cent in 2005 to 24.2 per cent in 2013 (OECD, 2016a). Several Latin-American countries have recently transitioned from two-shift schemes, where some grades are taught in the morning and some in the afternoon, to one-shift schemes that feature a longer school day (see Section 3.2). As other examples, President Obama in 2009 and Chancellor Osborne in 2016 advocated for longer school days in the US and UK respectively. In the US the National Center on Time and Learning (NCTL) promotes extended school schedules.

To address these questions, we take advantage of two attractive features of the Chilean educational system. Firstly, we exploit the passing of the Full School Day (*Jornada Escolar Completa*, or FSD henceforth) reform in 1997, which substantially increased daily instruction time in all publicly subsidized (*i.e.* public and charter) primary and secondary schools, whilst leaving the term length and the national curriculum unchanged. The increment was sizable, ranging from 4 to 9 additional instruction hours per week depending on the grade. In grades 1 to 4 it translated into a 26.7 per cent increase in weekly instruction time. Schools could decide when to adopt the longer school day and how to allocate the additional time across subjects. Secondly, we leverage the fact that in Chile there are public and charter schools. Both types of schools are funded through a voucher system, but differ substantially in terms of ownership and in the degree of autonomy that they enjoy: charter schools have more autonomy over staff decisions and over the course offer and content.

We estimate the effect of additional instruction time on achievement by exploiting within-school variation in years of exposure to the FSD across several cohorts of pupils who started grade 1 between 2002 and 2010, and who later took standardized reading and mathematics tests at the end of grade 4. As the availability of longer schedules may affect the composition of pupil intake, we restrict our attention to cohorts of *incumbent* students, *i.e.* those who started primary education in schools that had yet not adopted the FSD and who may have become exposed to it at different stages of their primary education. We further deal with potentially endogenous mobility across schools by instrumenting actual exposure with the exposure a student would have accumulated if he/she had remained in the school where he/she initially enrolled.

Our preferred linear specifications show that an additional year of exposure to the FSD raises fourth grade reading scores by 0.024 standard deviations (σ). The effect on mathematics scores is smaller (0.007-0.008σ) and non significant.

We then explore heterogeneity by pupils' socio-economic characteristics. The effect of additional instruction time on academic performance is smaller for students from advantaged backgrounds than for the rest of the pupils in our sample. This finding is consistent across different measures of household resources, which include parental education and the availability of books, a computer and a connection to the Internet at home. Although non negligible, the difference in the benefits of additional instruction time for pupils from different backgrounds is however not statistically significant. With this caveat about the precision of our estimates in mind, results suggest that returns to additional time at school are higher for pupils who - according to information drawn from time-use surveys - have scarcer support available outside of school. We further provide evidence suggesting that the longer school day is associated with a reduced frequency of homework. As this is likely to particularly benefit students with limited support outside of school, it could be a potential mechanism underlying these findings.

Then, exploiting the features of the Chilean setting, we study how school institutions affect returns to instruction time. In order to avoid capturing differences in the characteristics of students attending different types of schools, we compare public schools to charter schools that do not charge fees. While fee-charging charter schools serve pupils from more advantaged backgrounds, public and no-fee charter schools cater to schoolchildren of similarly lower socio-economic status. We document greater benefits of longer schedules for pupils who start primary education in no-fee charter schools than for those who start in public schools. The difference is large for both subjects - mathematics and reading - but is only statistically significant for the latter. Moreover, it does not decline when, in the same regression specification, we allow for the effect of longer schedules to differ depending on students' socio-economic background; this further suggests that the estimated difference does not reflect heterogeneity in the characteristics of students attending the two types of schools.

While survey evidence suggests that public and charter schools allocate the additional instruction time across subjects in a similar way, we uncover a significant difference in how public and no-fee charter schools adjust total contract and teaching hours to provide longer schedules: no-fee charter schools rely more on hiring new classroom teachers and less on increasing work hours per teacher than public schools, as shown by the evolution of the number of classroom teachers and their contract hours after the adoption of the FSD. It therefore appears that the higher degree of autonomy enjoyed by charter schools over staff decisions allows them to adjust the teaching input in a different way. Moreover, we show that public school teachers display a lower degree of satisfaction with the FSD scheme than their colleagues in charter schools. If extended teachers' working hours translate into a lower quality of time use at school, this could be one mechanism underlying the documented heterogeneity. The literature has found that charter schools are typically associated with better learning outcomes. Our findings suggest that in the context of a large-scale expansion of a school input (*i.e.* instruction time), charter schools may be able to adjust other inputs (*i.e.* teaching hours) in a more effective way.

The rest of this paper is organized as follows. Section 3.2 reviews the related literature. Section 3.3 describes the Chilean education system and the FSD reform. Section 3.4 presents the identification strategy. Section 3.5 describes the data and the sample. Section 3.6 discusses the main findings. Section 3.7 concludes.

3.2 Related Literature

This paper speaks to the growing literature that examines the effects of school institutions on pupils' performance. Several papers focus on newly founded or converted charter schools in the US. Studies on oversubscribed charter schools exploit the fact that admission depends on a lottery and document a positive effect both on school performance (Abdulkadiroğlu et al., 2011; Dobbie et al., 2011) and on medium-term non academic outcomes (Dobbie and Fryer Jr, 2015). Dobbie and Fryer Jr (2013) and Angrist et al. (2013) highlight that high-performing charter schools are characterized by certain practices and features, among which there is increased instruction time. Along this line, Baude et al. (2014) show that the quality of charter schools in Texas improved between 2001 and 2011, largely driven by the adherence to *No Excuses*-style curricula. Abdulkadiroğlu et al. (2016) analyze charter takeovers (*i.e.* formerly public schools converted into charter schools) and also report positive effects on achievement of grandfathered students. Eyles et al. (2017) and Eyles and Machin (2018) examine how the conversion of some English community schools into academies - autonomous, state-funded education establishments not subject to local authority control - affects achievement of legacy enrolled pupils. The former work studies several post-2010 episodes of conversion involving already high-performing primary schools and does not find significant effects on achievement. The second assesses the first round of conversion of mostly underperforming secondary schools in the 2000s, documenting instead a positive impact on test scores.

Our work is also related to the literature on the relationship between instruction time and achievement.[2] A set of papers studies the effect of the number of school days prior to standardized tests on performance, by exploiting either unplanned school closures due to adverse weather conditions (Marcotte, 2007; Marcotte and Hemelt, 2008; Hansen, 2011; Goodman, 2014) or changes in term dates and/or test dates (Sims, 2008; Agüero and Beleche, 2013; Aucejo and Romano, 2016). These studies find positive, although in some cases modest, effects. While they leverage small variations in the number of school days, we focus on substantial and permanent changes to the length of the school day. Varying the length rather than the number of school days may have different consequences on student achievement. For example, while the former entails a re-organization of daily routines, the latter does not.

Starting from Lavy (2015), recent studies examine the effect of instruction time on

[2]The early studies mostly focus on term length and report modestly positive to insignificant effects. These studies rely either on variation in term length between and within US states over time (Rizzuto and Wachtel, 1980; Card and Krueger, 1992; Grogger, 1996; Betts and Johnson, 1998; Eide and Showalter, 1998) or on cross-country differences (Lee and Barro, 2001; Woessmann, 2003). A review of studies conducted in the 1985-2009 period can be found in Patall et al. (2010).

achievement and the drivers of its effectiveness by using cross-country PISA data and exploiting within-pupil or within-school variation in subject-specific classroom hours. Lavy (2015) finds that a one-hour increase of weekly subject-specific instruction time raises scores by 0.06σ and that school and student characteristics matter: the effect is larger for schools that enjoy more autonomy and for pupils from disadvantaged backgrounds. Rivkin and Schiman (2015) further highlight that productivity of instruction time depends positively on the quality of the classroom environment, as captured by student disruption and student-teacher interactions. Cattaneo et al. (2017) focus their attention on Switzerland and document that students in more demanding school-tracks enjoy greater benefits. Also these studies leverage a source of variation different from ours. Different allocations of weekly instruction time across subjects do not necessarily entail a change in the length of the school day. Students do not have to re-arrange their daily routine or reduce the time for activities carried out outside of school, nor do schools need to operate for more hours.

A number of papers exploit instead reform-induced variation in instruction time. Pischke (2007) and Parinduri (2014) study the effects of exceptionally short or long school years due to country-level reforms of school calendars that leave the curriculum unchanged.[3] More similarly to us, Huebener et al. (2017) and Lavy (forthcoming) examine reforms that increase weekly instruction hours in Germany and Israel, respectively. They both find a positive effect on achievement. The former documents a larger gain for high-performing students, while the latter does not find evidence of differential benefits across pupils from different socio-economic backgrounds. Battistin and Meroni (2016) and Meroni and Abbiati (2016) study an expansion of mathematics and reading instruction time in lower secondary schools in southern Italy, documenting positive effects on mathematics test scores, concentrated among high-achieving disadvantaged pupils. Recently, Figlio et al. (2018) show that extending the school day and providing additional literacy instruction time in low-performing schools in Florida have a positive effect on reading test scores.

Similarly to Chile, several other Latin American countries have switched from a two-shift scheme - where some grades are taught in the morning and some in the afternoon - to a one-shift scheme, substantially lengthening the school day. The impact has been evaluated in a series of reports.[4] Findings are mixed, suggesting that

[3]The former studies the short 1966-67 West German school year and documents an increase in repetition rates in primary school as well as a reduction in enrollment to higher secondary school tracks, but no effects on earnings and employment. The latter examines the long 1978-79 Indonesian school year and reports a reduction in repetition rates and improved educational attainment, with positive effects also on wages and on the probability of working in the formal sector.

[4]Cerdan-Infantes and Vermeersch (2007) on Uruguay, Almeida et al. (2016) on Brazil and Hincapie (2016) on Colombia.

how the reform of school schedules is implemented and how additional instruction time is used play important roles in shaping returns.

Two papers study the effect of the FSD reform in Chile on standardized test scores. Bellei (2009) focuses on performance at grade 10 in 2001 and 2003, adopting a difference-in-differences approach. Berthelon et al. (2016) explore the effect on early literacy skills at grade 2. Based on one year of observations (2012), they instrument exposure to the FSD with the local availability of schools offering longer schedules. Both papers find positive and significant effects on performance. We focus on all cohorts that start primary school between 2002 and 2010 and examine the effect of the FSD on their academic performance in a different grade (grade 4). In addition, we propose a different identification strategy to assuage concerns about student endogenously sorting into schools offering the FSD. Furthermore, we examine how schools' types and students' characteristics affect returns to longer schedules and explore some mechanisms that could explain the heterogeneous effects. In a recent paper Dominguez and Ruffini (2018) study the effect of the FSD on longer-term outcomes; they focus on educational attainment and earnings when individuals are in their 20s and 30s, finding positive effects.

3.3 Institutional Setting

3.3.1 The Chilean School System

The Chilean school system features two education cycles: primary education (grades 1-8) and secondary education (grades 9-12). Standardized tests called SIMCE assess pupils' knowledge and skills in core subjects at the end of various grades. The testing frequency is highest in fourth grade, with tests taking place every year since 2005.[5]

Education is provided by three types of schools: public schools, charter schools and non subsidized private schools. Public schools are free and are funded through student vouchers.[6] Charter schools are private, but they are publicly subsidized through the voucher system as well. Since 1994 charter schools can also charge tuition fees, but the size of the voucher decreases as tuition fees increase. Non subsidized private schools are funded only through tuition fees and are usually substantially more expensive than charter schools.

The FSD reform applies to public and charter schools, which serve more than 90

[5]Fourth graders were also tested in 1999 and 2002.

[6]During the 1980s the Chilean school system experienced a major transformation that transferred the administration of public schools from the Ministry of Education to Municipalities. Furthermore, the funding system was changed by introducing a voucher that could be used both in public and charter schools.

per cent of the students attending regular programs in the school system.[7] Despite both being publicly subsidized, they are different in how they are managed and regulated. Public schools are either managed by the Municipal Department of Education (*DAEM*) or by non-profit Municipal corporations.[8] The working conditions are regulated by a labor code specific to education professions.[9] Charter schools are private organizations and, accordingly, the working conditions of teachers are regulated by the private sector labor code.[10] Different regulations translate into charter schools having greater autonomy and flexibility in the management of the teaching staff, in terms of recruiting, dismissal and compensation policies. They also enjoy more responsibility and freedom over the design of the curriculum. In Appendix Section 3.A we discuss in more detail the main regulatory differences between public and charter schools. We also provide further evidence when exploring the differential effect of the FSD by type of school in Subsection 3.6.3.

3.3.2 The FSD Reform

In 1997 the Chilean government decided to increase daily instruction time in all publicly subsidized primary and secondary schools (*i.e.* public schools and charter schools), whilst leaving the term length and the national curriculum unchanged.[11] The Full School Day (*Jornada Escolar Completa*, or FSD henceforth) reform aimed at improving the quality of education and reducing inequality in learning outcomes.[12] The reform envisaged a substantial increase of instruction hours (which last 45 minutes). Specifically, in primary schools 8 instruction hours were added in grades 1 to 6 and 5 hours in grades 7 and 8; in secondary schools, grades 9 and 10 experienced a 9-hour increase of instruction time per week.[13] In grades 1 to 4, this translated into a 26.7 per cent increase of weekly instruction time. As a result, in 2015 the length of school days in Chilean primary schools was the highest among OECD countries, when considering total compulsory instruction time (OECD, 2016b).

Schools could choose when to implement the FSD.[14] The deadline was initially set to 2002. However, it was later extended and differentiated by type of school and student: 2007 for all public schools and for charter schools catering to disadvan-

[7] This figure excludes education for adults, education for students with specific disabilities and other types of special programs.

[8] While the director of the *DAEM* is usually a teacher appointed by the Municipality, corporations are led by a board of directors who do not need to be teachers and whose president is the major of the Municipality.

[9] This is called *Estatuto de los Profesionales de la Educación*.

[10] This is called *Código del trabajo*.

[11] Increasing daily instruction time is not mandatory in grades 1 and 2.

[12] The Law 19494 that introduced the FSD was published on January 25, 1997.

[13] In grades 11 and 12, 6 instruction hours were added in the scientific-humanities track, while 4 hours were added in the vocational track.

[14] Schools could also adopt the FSD in different grades at different times, but they were mandated to ensure that pupils who started attending the longer school day in a given grade would then also be exposed in all following grades.

taged pupils, 2010 for the rest of charter schools. Yet, by 2013 - the last year of data available to us - there were still schools operating under the old scheme. Figure 3.1 illustrates the pattern of adoption of the FSD between 1997 and 2013 for primary schools. For every year, it shows the number of schools, as well as the share of public and charter schools, that had adopted the policy by that year. The two types of schools display similar patterns of adoption, although a larger share of public schools had implemented the FSD by 2013.[15]

By the time the reform was announced many schools were operating a two-shift scheme: some grades were taught in the morning and some in the afternoon. The increased instruction time and the longer school day required a change to a one-shift scheme, where all pupils attend school from the morning to mid-afternoon. Table 3.1 illustrates the daily school schedules with and without the FSD, inclusive of time devoted to breaks. Without the FSD pupils spend at least 5 hours per day at school. The typical morning shift runs from 8.00 to 13.00, while the typical afternoon shift runs from 14.00 to 19.00. Under the FSD students spend at least 7.08 hours per day at school. If the school adopts the FSD from Monday to Friday, the typical school day starts at 08.00 and ends at 15.05. If the school adopts the FSD on 4 days and the shorter school day on the remaining one, the typical longer school day starts at 8.00 and ends at 15.45.[16]

The passage from a two-shift to a single-shift scheme implies that pupils have lunch at school. For most students, however, this did not translate into a substantial change in the nutritive content of their diet. First, Chile had virtually no problem of infant malnutrition when the FSD reform was passed.[17] Second, students from disadvantaged backgrounds could also have lunch at school under the short school day scheme. The main difference between the short and the long school day is therefore the increase of instruction time, which requires adjusting the teaching input.

Table 3.2 reports weekly instruction hours per subject with and without the FSD for grades 1 to 4. It shows that the legislated increase in instruction time was not allotted to specific subjects, but rather allocated to the so-called "Free Choice time", which schools could decide how to use. Therefore, schools had considerable freedom over the organization of the FSD, the only constraint being the approval by the Ministry of Education (*Ministerio de Educación*) of a pedagogical plan that described the use of the additional time.

We do not observe how each school allocates the additional time across subjects. However, we can provide some evidence based on a survey administered in 2005 to

[15] By 2013 around 12 per cent of primary schools were still operating without the FSD.

[16] The minimum hours of daily instruction are prescribed by the law. Schools can freely choose the time at which the school day starts. The daily schedules in Table 3.1 are built assuming that the full school day and the morning shift start at 8.00, while the afternoon shift starts at 14.00.

[17] In 2000 only 2.9 per cent of children aged 0-6 suffered from malnutrition and only 0.3 per cent suffered from moderate or serious malnutrition (Mönckeberg B., 2003).

investigate the use of time in fifth grade at 387 urban primary schools that had already implemented the FSD at that point.[18] Drawing on this, Table 3.3 reports the allocation of weekly instruction time across curricular subjects, both for all schools (columns 1-2) and distinguishing between public (columns 3-4) and charter (columns 5-6) schools. "Core time" excludes "Free Choice time". It shows that schools devote a substantial portion of "Free Choice time" to core subjects. Among those, more hours are allocated to Spanish than to mathematics.[19] A small fraction of additional instruction time is dedicated to other subjects; the remaining portion of "Free Choice time" is distributed among various extra-curricular activities (not reported in the table for brevity). Charter schools devote slightly less additional time to Spanish and mathematics. However, the allocation of additional instruction time across subjects is similar in public and charter schools. The only significant differences emerge with regards to foreign languages and religion, to which charter schools devote more of the additional instruction time.

To further investigate the effect of the FSD on schools' time use, we rely on data that reports, at the school-class-year level, the list of subjects taught.[20] Appendix Figure 3.C.1 shows, in an event study framework that collapses information at the school-grade-year level, the evolution of the total number of subjects, as well as subjects related to specific disciplinary areas, around the adoption of the FSD.[21] Following the introduction of the longer school day in a given grade, the number of subjects taught in that grade increase by a small but statistically significant amount, up to almost 0.1 four years after the implementation of the policy. The increase is driven by the fact that there are more subjects related to foreign languages as well as more tutorials and workshops. While the number of subjects is an imperfect measure of how schools use the additional time, as they could simply increase the hours devoted to each subject, this provides further evidence that the longer school day translated into an increase of instruction time.

Augmenting instruction time and lengthening the school day generated additional

[18]The survey was administered by the Studies Directorate of the Sociology Faculty at the Catholic University of Chile (*DESUC*) and a report based on it was written by Ruz Pérez and Madrid Valenzuela (2005).

[19]Spanish also features more instruction time also under the shorter school day.

[20]This data is available at `http://datos.mineduc.cl/dashboards/19923/bases-de-datos-de-planes-y-programas-de-estudios-anos-2002-al-2016/`.

[21]The event study specification reads:

$$Y_{gst} = \gamma_g + \eta_s + \theta_t + \lambda_{sg} + \phi_{st} + \mu_{gt} + \sum_{\rho=-5}^{-2} \beta^{\rho} \mathbb{1}(p_{gst} = \rho) + \sum_{\rho=0}^{4} \beta^{\rho} \mathbb{1}(p_{gst} = \rho) + \varepsilon_{gst} \tag{3.1}$$

g, s and t index the grade, the school and the year, respectively. $p_{gst} = t - E_{gs}$ is the distance (in years) from the event, which is the introduction of the FSD in grade g of school s. Controls consist of grade (γ_g), school (η_s) and year (θ_t) fixed effects, as well as their interactions. The FSD is adopted in event-year 0 and coefficients β^{ρ} show how different the number of subjects taught is in event-year ρ relative to event-year -1, which is taken as the reference year. Standard errors are clustered at the school-grade level.

operational costs, which were funded through an increase in the baseline vouchers, by 25-50 per cent depending on the grade (Table 3.2).[22] Some schools also had to expand their infrastructure in order to switch to the single-shift scheme. Infrastructure-related costs were funded through *ad-hoc* additional resources, which were allocated through public tenders organized by the Ministry of Education and its regional offices. Priority was usually granted to schools catering for students from lower socio-economic backgrounds.[23]

3.4 Empirical Strategy

In order to study whether increased instruction time and a longer school day affect achievement, we exploit the fact that we observe several cohorts of pupils starting primary education in a given school in the 2002-2010 period and then taking a standardized test at the end of grade 4 – possibly in a different establishment – over the period 2005-2013. Since we can follow the entire school career of each one of such students, we can compute actual years of exposure to the FSD by the end of grade 4 as $ExpFSD4 = \sum_{j=1}^{4} d^j$, where d^j is a dummy that takes value 1 if the pupil is exposed during grade j to the FSD.[24]

We then estimate the following specification:

$$Y_{ist} = \eta_s + \theta_t + \beta ExpFSD4_{ist} + \gamma X_{ist} + \delta Z_{st} + \varepsilon_{ist} \tag{3.2}$$

Y_{ist} is the test score of student i who starts primary school in school s in year t and then takes the standardized test at the end of grade 4. η_s is a set of school fixed effects that account for time-invariant heterogeneity across schools; θ_t is a set of year fixed effects that control for common unobserved year-specific shocks. In the richest specifications, we also include a set of controls at the student and at the school level. Specifically, X_{ist} is a vector of student characteristics measured in first grade, which include: gender, age, attendance rate and end-of-year status (*i.e.* promotion to second grade or retention in first grade). Z_{st} averages student characteristics contained in X_{ist} at the school level. It also includes enrollment and

[22] The final amount that a school receives through student vouchers also depends on its location, size, and other characteristics. We report the increase of the baseline voucher, because this was the change common to all schools.

[23] Yet schools serving pupils from higher socio-economic backgrounds were less likely to need infrastructure-related investments.

[24] The treatment is therefore more precisely defined as the number of grades attended at least once under the FSD scheme by the end of grade 4. Furthermore, exposure during grade 4, the grade when students take the standardized test, is computed taking into account only the first year when a pupil attends grade 4. This is motivated by the choice of considering only the first observation when schoolchildren belonging to the master sample (see Section 3.5) take the test multiple times. Throughout the paper, we use the term years of exposure to the FSD for brevity. Moreover, the two definitions are exactly equivalent for non repeaters, who constitute the largest majority of the sample (88 per cent).

average class size in first grade. Standard errors are clustered at the school level.

By including first-grade school fixed effects, specification (3.2) leverages variation in exposure to the FSD by the end of grade 4 across cohorts of students who enrolled in the same establishment. It therefore exploits the fact that, depending on whether the school adopted longer schedules within our sample period (*i.e.* by 2013) and on the exact year of adoption, adjacent cohorts of enrollees could experience a different exposure to the FSD before taking the test. This source of variation can be used to estimate the causal effect of the FSD on learning outcomes if cohorts of pupils are not systematically different along characteristics that are not taken into account in specification (3.2) and that correlate both with years of exposure to increased instruction time and with achievement.

Given the staggered adoption of the FSD across schools, a first concern is that parents would factor the availability of the longer school day into their preferences about the school in which to enroll their children. This could affect the composition of pupil intake, possibly along dimensions that our set of controls cannot account for. According to parent surveys administered alongside the test in 2005, the FSD was the most important reason for enrolling their child in a given school for only 2.46 per cent of parents. Proximity to home (27.61 per cent), the presence of a relative in the school (15.66 per cent) and the school's prestige (18.89 per cent) were cited as the most important determinants of school choices.[25]

Nonetheless, we address this concern by restricting our analysis to *incumbent* pupils. This means that we only consider pupils who enroll in first grade in a given school before such school adopts the FSD. As an example, if a school adopts the longer school day in 2007, we discard students who start primary education in that school in 2007 or later. Cohorts who enrolled before 2007, on the other hand, made their decision before the introduction of the FSD and possibly became exposed to it at some point in their school career. For *incumbent* students who never repeat a grade the range of the treatment variable ($ExpFSD4$) is 0-3, as exposure to the FSD cannot start earlier than in grade 2. For repeaters, on the other hand, the range is 0-4; the variable takes value 4 in cases when the school adopts the FSD in the year when the pupil is repeating first grade. Furthermore, restricting the sample to *incumbent* students implies that first-grade controls included in specification (3.2) are observed before the adoption of the FSD and hence are pre-determined with respect to the treatment. The focus on legacy enrollment cohorts also characterizes recent studies on the effects of charter takeovers in the US and of academy conversions in England (Abdulkadiroğlu et al., 2016; Eyles et al., 2017; Eyles and Machin, 2018). This restriction attenuates identification issues related to unobserved changes in pupil intake, the more so the less parents can anticipate the exact year in which a school

[25]The ethical values (8.47 per cent) and the cost (7.34 per cent) of the school follow in the ranking. The presence of the FSD is ranked eight among fifteen options.

will adopt the FSD.

Students can move across schools and in Chile school transfers are indeed a common phenomenon; in our master sample (described in Section 3.5) around 35 per cent of students change school between grades 1 and 4. Pupils who enroll in first grade in the same establishment and in the same year can therefore experience a different exposure to the FSD by the time they reach grade 4. Furthermore, if mobility across schools is influenced by the availability of longer schedules, student-level actual exposure to the FSD could be correlated with other unobserved determinants of achievement. To mitigate this concern, in our preferred specification we instrument actual exposure to the FSD with the exposure a student would have accumulated had he/she never transferred from the school where he/she attended first grade. The instrumental variable is therefore $PotExpFSD4 = \sum_{j=1}^{4} d_s^j$, where d_s^j is a dummy variable that takes value 1 if the student would have ever been exposed to the FSD in grade j, had he/she remained in school s, where he/she started first grade.[26]

When discussing results in Section 3.6, we will show that the instrument is relevant, as there is a positive and statistically significant relationship between "potential" exposure and actual exposure to the FSD by grade 4. By relying on this instrument we aim to isolate and exploit the variation in actual exposure that is not affected by possibly endogenous mobility decisions of *incumbent* pupils after first grade. Moreover, we assume that "potential" exposure is not systematically correlated with unobserved determinants of achievement and affects fourth grade test scores only through its impact on actual exposure.

A remaining concern is that the timing of adoption may depend on past performance. For example, if schools switch to the longer school day after they observe a cohort of pupils faring particularly poorly at the test, our estimates may simply capture mean-reversion effects. In general, there can be concerns about confounding effects of underlying school-specific trends in test scores. We show in Section 3.5 that there are no visible clear trends in reading and mathematics scores in the years preceding the switch to longer schedules. Another concern is that other events may take place at the school around the time of FSD adoption, which could also affect learning outcomes in the following years. We discuss and address these further issues in Appendix Section 3.B, where we show that findings remain similar when: (i) we restrict our attention to pupils starting primary school in establishments that did not receive public funds for expanding infrastructure; (ii) we control for another policy

[26] To build this "potential" measure of exposure we also assume that the student would have never repeated, as we do not observe the pattern of repetitions in this counterfactual school career. Therefore, the range of the instrumental variable for all *incumbent* pupils is 0-3. The first stage regression specification reads:

$$ExpFSD4_{ist} = \eta_{0,s} + \theta_{0,t} + \beta_0 PotExpFSD4_{ist} + \gamma_0 X_{ist} + \delta_0 Z_{st} + \epsilon_{ist} \tag{3.3}$$

targeting disadvantaged students implemented in 2008 or we focus our attention only on cohorts of students not exposed to such a policy.

3.5 Data and sample

We link several administrative and survey datasets on account of unique school, student and teacher identifiers.

Data on achievement in fourth grade comes from a nationwide standardized low-stakes test (SIMCE test) designed by the Education Quality Agency (*Agencia de Calidad de la Educación*).[27] It is administered at the end of the school year and is marked by external examiners, therefore leaving little room for test score manipulation. Individual records on performance in the test are available for fourth grade students in 1999, 2002 and then with a yearly frequency from 2005 onward. We restrict our attention to the 2005-2013 waves of the test. The reason is that we can follow students' school careers only for cohorts who start primary school from 2002 onward; this is necessary both to correctly identify *incumbent* students (*i.e.* pupils who enroll in first grade in a school that has not yet adopted the FSD) and to compute actual exposure to the FSD for students who move across schools between grades 1 and 4. 2013 is the last year of data available to us. We use pupil-level test scores in the reading and mathematics sections of the test as our measure of achievement. Scores are standardized by year and subject to have a mean equal to 0 and a standard deviation equal to 1. Alongside the test, surveys are administered to students and their parents, as well as to teachers. Based on questions that are consistent across all waves of the parent survey, we recover a rich set of information on pupils' backgrounds as of grade 4 that we use to study heterogeneity by students characteristics. Based on teacher surveys, we provide evidence on the frequency of homework assignments in schools with and without the FSD.

The second source of information is the register of pupils enrolled in the school system over the period 2002-2013, which is maintained by the Ministry of Education. Besides gender and date of birth, for every school year it records information about the school that the student attends, the attendance rate and the end-of-year status (*i.e.* promotion to the next grade or retention in the same grade). We also have access to the register of educational establishments, from which we recover the location (an urban or rural area) and the administrative status of the school (public, charter or non subsidized private). A companion dataset records the year of adoption of the FSD at the school-grade level over the period 1997-2013. Based on these sources, we reconstruct the school career from grade 1 to grade 4 of every student who started primary school between 2002 and 2010 and took the fourth grade test

[27] While the stakes are low because the test does not impact a student' final evaluation, schools care about it because school-level average scores are publicly available for consultation.

between 2005 and 2013; we then compute the actual years of exposure to the FSD by the end of grade 4, as well as the exposure a student would have experienced had he/she never transferred from his/her first grade school. We also retrieve the set of first grade student- and school-level characteristics that we include in the richest regression specification. In order to distinguish charter schools with and without tuition fees, we rely on a dataset maintained by the Ministry of Education that records all the subsidies that schools received from the government over the 2005-2013 period. Since charter schools that charge tuition fees receive reduced subsidies, we can distinguish them from schools that do not charge tuition fees.[28]

We also exploit the information contained in the register of teachers, which is available for the period 2003-2013. We draw on this dataset to study how no-fee charter and public schools adjust the number of teachers and their working hours after the adoption of the FSD. We also rely on the 2005 Longitudinal Teachers Survey (*Encuesta Longitudinal Docente*) to investigate differences in teachers' opinions on the FSD;[29] based on the 2015 Time-Use Survey (*Encuesta Nacional de Uso de Tiempo*), we examine if academic support outside of school varies across pupils from different socio-economic backgrounds.[30]

Finally, we digitized from primary sources the list of schools that received additional funds to expand their infrastructure when lengthening the school day; we parsed the releases of the Official Journal (*Diario Oficial*) published by the Interior Ministry (*Ministerio del Interior*) over the period 1997-2004 and searched for the outcomes of all public tenders through which *ad-hoc* resources for infrastructures were assigned.[31] Based on this, we create a dataset that records, for every school, the year in which resources were disbursed and the amount received, if any. Since 2008, students from disadvantaged backgrounds are granted additional subsidies (PSS) on top of the vouchers. We obtain the list of beneficiaries from the Ministry of Education. This information is used to perform robustness checks described in Appendix Section 3.B.

In order to create the master sample of our analysis, we restrict our attention to *incumbent* pupils, *i.e.* students who started first grade in a given school when the FSD had not been yet introduced (see Section 3.4). Furthermore, we discard students that attended non subsidized private schools at some point between grades 1 and 4. This is motivated by the fact that the FSD reform only applies to publicly subsidized schools. Moreover, we do not know whether a given non subsidized private school was already offering a longer school day or started providing it at some

[28] We classify a charter school as a no-fee charter school if it never charged fees between 2005 and 2013. For 2.39 per cent of charter schools attended in first grade by students belonging to the master sample we do not find information about the tuition fees in the dataset.

[29] The Longitudinal Teachers Survey was implemented by the Microdata Center of the University of Chile.

[30] The database of the 2015 Time-Use Survey can be downloaded from `https://www.ine.cl/estadisticas/menu-sociales/enut`.

[31] The last tender took place in 2004.

point after it became compulsory for other types of schools. Therefore, students attending non subsidized private schools cannot serve as a control group.

The master sample consists of around 600,000 fourth-grade test takers; they started primary school between 2002 and 2010 in schools that had not yet adopted the FSD and took the test between 2005 and 2013. It follows that schools attended by pupils in the master sample had not switched to the longer school day by 2002. Given that the first transitions to the single-shift occurred in 1997, our sample of schools consists of mid to late adopters.

As discussed in Section 3.4, a threat to identification could arise if schools adopted the FSD based on the trend or transitory component of test scores. Figure 3.2 plots coefficients from an event study exercise where the specification reads:

$$Y_{ist} = \eta_s + \theta_t + \sum_{\rho=-5}^{-2} \beta^{\rho} \mathbb{1}(p_{st} = \rho) + \sum_{\rho=0}^{4} \beta^{\rho} \mathbb{1}(p_{st} = \rho) + \varepsilon_{ist} \tag{3.4}$$

Y_{ist} is the reading or mathematics score of student i who takes the SIMCE test in school s in year t. $p_{st} = t - E_s$ is the distance (in years) from the event, which is the introduction of the FSD in at least one grade in school s. η_s and θ_t are school and calendar year fixed effects, respectively. The FSD is implemented in event-year 0 and coefficients β^{ρ} show how different scores in event-year ρ are relative to event-year -1, which is taken as the reference year. Schools are observed up to 5 years before and after the introduction of the longer school day and the sample consists of all schools where students in the master sample enrolled in first grade.[32] For both subjects, there appear not to be evident trends in the pre-adoption period, suggesting that test scores were not trending either downward or upward before schools decided to implement longer daily schedules. Furthermore, there are no evident spikes or dips in test scores just before the introduction of the FSD. On the other hand, from event-year 1 scores start increasing, suggesting a positive effect of the FSD on achievement. We will then provide a formal estimation based on our identification strategy in Section 3.6.

Table 3.4 reports summary statistics for pupils in the master sample. Column (1) pools all students together, whereas columns (2) to (4) split schoolchildren according to the type of school (public, charter without tuition fees or charter with tuition fees) they attended in first grade. In the vast majority of households (87 per cent) parents do not have university education.[33] Only 15 per cent of students have more than

[32] The sample is unbalanced, meaning that not all schools are observed in every event-year. Given the calendar of SIMCE tests, using a balanced sample would significantly reduce the number of event-years that we can observe. For this exercise, we also use the 1999 and 2002 waves of the SIMCE test.

[33] We build a variable that measures parental education by setting $ParentalEd = max(MotherEd, FatherEd)$, where $MotherEd$ and $FatherEd$ are the highest mother and father's academic attainment, respectively; if the information for either one of the two parents is

50 books at home; 55 per cent of the households have a computer at home and slightly less than one third also have a connection to the Internet.[34] The first grade attendance rate is very high (94 per cent) and 3 per cent of pupils repeat first grade. On average, there are 35 students in a first grade class.

When splitting students according to the type of establishment they started primary school in, it emerges that public schools and charter schools without tuition fees cater to relatively similar students. On the other hand, schoolchildren attending charter schools that charge tuition fees live in more affluent households. Test scores are lowest in public schools and highest in charter schools with tuition fees.

3.6 Results

In this section we first discuss the average effect of the FSD on all pupils in the master sample (Subsection 3.6.1). We then explore whether and how effects are heterogeneous depending on the resources available in the household where the student lives (Subsection 3.6.2) and on the type of school he/she attends (Subsection 3.6.3). Heterogeneous effects of additional instruction time are interesting to study because increasing the amount of time that pupils spend at school reduces the amount of time they can devote to other activities outside of school. The return to longer school schedules therefore could depend on the absolute quality of time use at school, which can vary across schools, and on its relative quality with respect to time use outside of school, which can vary among students.

When investigating heterogeneous effects, we estimate a richer version of specification (3.2), whereby we also interact the treatment and all controls with a dummy variable D that captures a given dimension of heterogeneity.[35] We report estimates coming from the preferred linear IV specification that includes the full set of controls (which we denote as FE-IV2).[36]

3.6.1 The effect of the FSD on achievement

Table 3.5 reports results from regression specification (3.2). We start by discussing coefficients when we estimate the most parsimonious specification, which

missing, we rely on the level of education achieved by the other parent.

[34]Information about students' backgrounds, i.e. parental education and resources at home, is drawn from parent surveys. Since these variables are observed at the end of grade 4, they could be affected by a student's exposure to the FSD (for example, if longer school days have an effect on parents' labour supply); for this reason, they are not included in the regression specifications, which only feature pre-determined controls. Furthermore this information is missing for around 15 per cent of schoolchildren in the sample.

[35]A fully interacted specification yields estimates that are equivalent to those obtained from separately estimating two regressions on the sub-sample where $D = 0$ and on the sub-sample where $D = 1$.

[36]$ExpFSD4$ and $ExpFSD4 \cdot D$ are instrumented with $PotExpFSD4$ and $PotExpFSD4 \cdot D$, respectively.

only includes school and year fixed effects, and we do not instrument actual years of exposure to the FSD (specification FE1, column 1). These estimates point to a virtually null effect on reading and a negative impact on mathematics. Including pre-determined controls listed in section 3.4, however, changes the picture significantly (specification FE2, column 2): the effect of an additional year of exposure to the FSD is positive for both subjects, although it is only statistically significant for reading (0.011σ). This indicates that controlling for first-grade status (pass or repeat) is important because repeaters, who are low performers, spend more years at school and are therefore more likely to be exposed to the FSD at some point.

As mentioned in section 3.4, a non negligible fraction of students transfer from one school to another between grades 1 and 4. Furthermore, the availability of longer daily schedules appears to influence mobility across schools. Appendix Figure 3.C.2 shows the evolution of transfers of pupils attending grades 1 to 4 at the school-year level, in a 5-year window around the implementation of the FSD. The estimated event study specification reads:

$$Y_{st} = \eta_s + \theta_t + \sum_{\rho=-5}^{-2} \beta^p \mathbb{1}(p_{st} = \rho) + \sum_{\rho=0}^{4} \beta^p \mathbb{1}(p_{st} = \rho) + \varepsilon_{st} \tag{3.5}$$

Following the introduction of longer schedules schools experience a decline in the outflow of pupils; at the same time, although the pre-adoption pattern is more scattered, inflows of students appear to increase, with a spike in the year of adoption. As a result, net transfers (*i.e.* the difference between transfers into and transfers out of a given school) grow, by up to 5 pupils per year. Appendix Table 3.C.1 further shows that, among schoolchildren who belong to the master sample, those who transfer are negatively selected, as they have a slightly lower attendance rate in grade 1 (93 per cent versus 95 per cent) and are more than twice as likely to repeat first grade.[37] Moreover, it emerges that pupils tend to transfer towards schools that offer the FSD; while transferring and non transferring students have a very similar "potential" exposure (*i.e.* the exposure they would have experienced had they remained in their first-grade schools), the former end up with a much higher actual exposure.[38] Partly because of fewer transfers out and more transfers in, the number of students per class in grades 1 to 4 increases after the adoption of the FSD (Appendix Figure 3.C.3) by an amount that is however modest (less than 1.50 additional pupils) when compared to the average class size in primary schools.

These patterns motivate the decision to instrument actual exposure ($ExpFSD4$) with the exposure a student would have experienced had he/she never transferred

[37] First-grade attendance rates have a very low dispersion, so that a 1 percentage point difference amounts to almost one fifth of a standard deviation.

[38] This also holds true also when restricting the comparison to students who never repeat between grade 1 and grade 4.

($PotExpFSD4$). When adopting the IV approach, estimates are remarkably stable across the most parsimonious specification (FE-IV1, column 3) and the specification featuring all controls (FE-IV2, column 4). An additional year of exposure to the FSD significantly raises reading test scores by 0.024σ. The effect on mathematics test scores lies in the narrow range 0.007-0.008σ, but is not statistically significant. As shown in the same table, the instrumental variable displays a positive and strong relationship with the treatment, as the first stage coefficient is statistically significant and equal to 0.72, implying that for slightly less than 30 per cent of pupils the real exposure and the "potential" exposure do not coincide.

In columns (5) and (6) of Table 3.5, we relax the assumption that every additional year of exposure has the same effect on achievement. We estimate the preferred IV specification in a fully non parametric way, by introducing a set of dummies for every possible level of exposure to the FSD and setting 0 years of exposure as the reference category.[39] The non parametric specification reveals that the effect of longer schedules increases more than linearly with exposure. Three years of exposure are associated with a 0.114-0.116σ increase of reading test scores, significant at the 1 per cent level, and a 0.057-0.058σ increase of mathematics test scores, significant at the 5 per cent level.

The IV estimates therefore show that the FSD has a positive effect on learning outcomes, which increases more than linearly with exposure and is stronger for reading than for mathematics. The stronger impact on reading may depend on the fact that a larger fraction of additional instruction time is devoted to Spanish than to mathematics (Table 3.3). The pattern of coefficients in the fully non parametric specification is consistent with added instruction time in earlier grades having a positive effect on achievement in later grades. Moreover, as the passage from a two-shift to a one-shift scheme implies a re-organization of daily routines, it may also be explained by the presence of adaptation costs that eventually fade away over time.

A possible remaining concern is that other events may happen in a school around the adoption of the FSD and affect learning outcomes in the following years. In Appendix section 3.B we show that our estimates are robust to: (i) restricting the attention to pupils who started primary schools in establishments that most likely did not expand their infrastructure at the same time when the FSD was adopted; (ii)

[39]The non-parametric specification therefore reads:

$$Y_{ist} = \eta_s + \theta_t + \sum_{k=1}^{2} \beta_k \mathbb{1}(ExpFSD4_{ist} = k) + \beta_3 \mathbb{1}(ExpFSD4_{ist} >= 3) + \gamma X_{ist} + \delta Z_{st} + \varepsilon_{ist} \quad (3.6)$$

Specification (3.6) highlights that we collapse 3 and 4 years of actual exposure into a unique category, as only very few pupils (*i.e.* students who repeat first grade in the year when the school adopts the longer schedules) attend all 4 grades under the FSD scheme. In the IV specification, the set of dummies that capture every possible level of actual exposure to the FSD are instrumented by a set of dummies that capture every possible level of exposure a student would have experienced had he/she never transferred out of his/her first grade school.

controlling for a policy granting further subsidies to disadvantaged schoolchildren since 2008.

3.6.2 Heterogeneity by students' backgrounds

In this subsection we explore whether the effect of the FSD varies depending on the characteristics of the environment students are exposed to when they are not in school. We focus our analysis on the role of household resources, as reflected by parental education and the availability of books and ICT technologies at home.[40] We rely on this information to distinguish schoolchildren from a more privileged background (for whom $D = 0$) from others (for whom $D = 1$).

Table 3.6 shows that longer schedules appear not to benefit in a significant way pupils from more advantaged backgrounds. An additional year of exposure to the FSD does not raise by a statistically significant amount reading and mathematics scores for children living in households where at least one parent has some university education (columns 1 and 4), there are more than 50 books (columns 2 and 5), or both a computer and a connection to the Internet are available (columns 3 and 6). On the other hand, reading scores increase by a significant amount for pupils living in households where neither parent has any university education (0.022σ), there are at most 50 books at home (0.022σ), and either a computer or a connection to the Internet is not available (0.024σ).[41] Also mathematics scores increase by a larger amount, which however never becomes significantly different from 0.

It has to be noted, however, that the documented difference, as captured by the interaction term $ExpFSD4 \times D$, although large in size, is not statistically significant. With this caveat concerning the precision of the estimates in mind, the analysis provides suggestive evidence that returns to an additional hour of instruction time tend to be larger for students who have fewer resources and opportunities available at home and for whom, therefore, the relative quality of time spent at school is higher.

Drawing on information coming from the 2015 Chilean Time-Use Survey, Appendix Table 3.C.2 shows that pupils from privileged backgrounds indeed receive more support outside of school. We restrict our attention to households where there is at least one child aged 5-18 and we divide them into two groups, depending on whether either the head of the household or the head's spouse has any university education ($U_{hh} = 1$) or not ($U_{hh} = 0$). In households where $U_{hh} = 1$, the percentage of heads of household and heads' spouses who declare that they help their children with their

[40] This information is drawn from parent surveys administered alongside the test. The non-response rate is similar across the variables and is around 15 per cent. This explains the smaller sample size.

[41] These figures are the sum of coefficients related to the main term $ExpFSD4$ and the interaction term $ExpFSD4 \times D$. They are significant at the 1 per cent level.

homework is 48 per cent, whereas this percentage drops to 33 per cent in households where $U_{hh} = 0$ (column 1). Summing up the minutes that they dedicate to helping with homework on a given day of the working week and on a given day at the weekend, there is a 15-minute difference in favor of households where $U_{hh} = 1$ (column 2). Assuming a uniform distribution of help across the days of the week, this would translate into a difference of around 50 minutes per week. It is also interesting to look at support by other providers, in the form of tutoring outside of school. 5 per cent of pupils aged 12-18 and living in households where $U_{hh} = 0$ receive some tutoring, as opposed to 12 per cent of students living in households where $U_{hh} = 1$ (column 3). In terms of minutes per day, the former receive tutoring for less than half the time than the latter (column 4).

Appendix Table 3.C.3 draws information about the frequency of mathematics homework from the teacher surveys administered alongside the SIMCE test in 2011, 2012 and 2013. The limited period for which this information is available does not allow us to study the evolution of homework's frequency around the adoption of the FSD in an event study framework such as the one in (3.5). Panel A considers all schools and shows that the frequency of homework is lower in schools with the FSD than in establishments without it. For example, the percentage of teachers assigning homework after every class is roughly 20 per cent in schools where the FSD is not in place, while it drops to about 12 per cent in schools that feature longer schedules. In panel B we restrict our attention to schools that had not adopted longer schedules by 2011 and we compare the frequency of homework between the years 2011 and 2013. In establishments that did not adopted the FSD in 2012 or 2013 (column 2), the frequency is very similar in the the two years. On the other hand, homework is assigned much less frequently in 2013 than in 2011 in establishments that switched to the FSD by that year (column 3).[42] Overall, there is therefore suggestive evidence that longer school schedules are associated with less homework. If the productivity of homework is higher for schoolchildren from advantaged backgrounds, because they have more support at home, the reduction of its frequency that seems to be associated with longer school schedules could be one of the mechanisms that explains the documented heterogeneity.

3.6.3 Heterogeneity by school type

The absolute quality of time use is likely to be the primary driver of additional instruction time's effectiveness. It is therefore important to study the contribution of school characteristics in shaping returns to longer schedules. The Chilean school system provides an attractive setting because public and charter schools, whilst

[42] As an example, in 2011 around 52 per cent of teachers working in schools that had not adopted the policy declared that they had assigned homework after almost every class. This figure remained the same in schools that had not adopted the FSD by 2013, while it fell to 31.82 per cent in schools that adopted it in 2012 or 2013.

being both publicly subsidized, differ in terms of ownership and in the degree of autonomy. As explained in Subsection 3.3.1, charter schools have more autonomy than public schools over the management of school resources and the design of the curricula. Appendix Table 3.C.4 reports answers to school surveys administered alongside the 2006 and 2009 waves of PISA tests, which ask about the tasks over which the principal, the teachers or the governing body of the school have considerable responsibility. The sample consists of all public and charter schools that offer primary education.[43] It confirms that principals as well as the teachers and the governing bodies of charter schools have greater autonomy in designing the curricula, as they can decide the course offer and the course content more frequently. Moreover, they are more likely to be responsible for the budget formulation and allocation. They are also in charge of taking personnel decisions, in terms of recruitment, promotions and dismissals.

We therefore study whether the FSD has a differential effect in public and no-fee charter schools. To this end, we compare pupils who attended a public school in grade 1 (for whom $D = 0$) to students who enrolled in a charter school that does not charge tuition fees in grade 1 (for whom $D = 1$). The choice of focusing on these two types of schools is motivated by the fact that, as shown in Table 3.4, charter schools that charge tuition fees cater to more affluent pupils, whereas public schools and charter schools without tuition fees serve pupils of similarly lower socio-economic status. As we aim to uncover the role of school institutions, we do not want to capture differences related to students' characteristics.

Table 3.7 shows that returns to additional instruction time are higher for students starting primary school in no-fee charter schools than for those who enrolled in public schools. The difference, as captured by the interaction term $ExpFSD4 \times D$, is sizable for both subjects and statistically significant with regards to reading (columns 1 and 5). The effect of an additional year of exposure to the FSD on reading test scores is more than three times larger for students starting primary school in no-fee charter schools (0.061σ) as opposed to in public schools (0.019σ). Mathematics scores are raised by a statistically insignificant 0.012σ for pupils attending grade 1 in public schools; the coefficient more than doubles to 0.029σ for enrollees of no-fee charter schools, although the associated p-value is slightly above 0.1. Pupils attending public schools and no-fee charter schools are similar. If anything, those attending the latter type of school are slightly more affluent. In Subsection 3.6.2 we provided suggestive evidence that longer schedules are less beneficial to students from advantaged backgrounds. Hence, small differences in student characteristics across the two types of institutions should not be responsible for the large observed differences in returns to longer schedules. In columns (2) to (8) we further show

[43] PISA tests are administered to pupils aged 15. We therefore restrict our attention to secondary schools that also offer primary education, which explains the very small sample size.

that the documented differential effect remains remarkably stable when we add the interaction between exposure to the FSD and a dummy capturing parental education (columns 2 and 6), the availability of books at home (columns 3 and 7), and the availability of ICT technologies at home (columns 4 and 8).[44]

As discussed earlier, charter schools have more autonomy over the design of the curriculum and personnel decisions. According to survey evidence provided in Subsection 3.3.2, public and charter schools allocate the additional instruction time across subjects in a similar way. Therefore, we turn our attention to personnel decisions and investigate how the adoption of the FSD affects the number of classroom teachers, and their contract and teaching hours using the event study specification outlined in (3.5). Consistent with the need to provide more instruction hours, Figure 3.3 shows that total classroom teachers' contract hours and teaching hours increase after the adoption of the FSD (top panel). Moreover, the pattern of coefficients is similar across public and no-fee charter schools and confidence intervals overlap. Indeed, Appendix Table 3.C.5 shows that the difference between public and no-fee charter schools is significant only in the last event-year. When total contract and teaching hours are divided by the number of classes (bottom panel), the differences between the two types of schools are never significant.[45]

An increase in the number of total contract and teaching hours can be achieved by adjusting both the number of classroom teachers and the number of contract/teaching hours per teacher. Figure 3.4 shows that the number of classroom teachers grow both in public and no-fee charter schools, but the increase is higher in the latter. On the contrary, contract hours per teacher increase more in public schools. Appendix Table 3.C.5 shows that these differences are statistically significant in most post-adoption event-years. Also teaching hours per teacher grow more in no-fee charter schools than in public establishments, with the difference been slightly smaller. Therefore no-fee charter schools rely more than public schools on expanding the number of teachers, whereas they resort less to increasing teachers' workload. It seems that autonomy over staff decisions allows charter schools to adjust the teaching input in a different way.

Appendix Table 3.C.6 reports teachers' opinions in 2005 about the FSD, dividing them according to the type of school (public or charter) in which they teach. Public

[44]In columns (2) to (8) all controls are also interacted with the dummy capturing a given characteristic of the household in which the student lives. Furthermore, specifications are estimated on the sub-sample of pupils for whom all background characteristics (parental education, number of books at home, availability of ICT technologies at home) are non missing. The number of observations is nonetheless slightly different across specifications because of differences in the number of singletons that are dropped.

[45]The coefficients presented in Table 3.C.5 come from a richer version of specification (3.5), where event year dummies and calendar year fixed effects are also interacted with a dummy (D_s) taking value 1 if the school is a no-fee charter school. This specification allows us to test whether differences between public and charter schools are statistically significant.

school teachers display a lower degree of satisfaction with longer daily schedules. Only 45 per cent of them judge the FSD as "good or very good", compared to 54 per cent of charter school teachers. This may signal that the workload of teachers in public schools increases excessively following the introduction of the FSD. This could in turn negatively affect the absolute quality of additional time use in public schools, contributing to explain the lower returns to longer schedules.[46]

3.7 Conclusions

Scholars and policymakers have long been interested in understanding the determinants of school effectiveness. Many countries undertake costly educational reforms in order to expand the resources available to schools. Some countries, including US and UK, have also implemented reforms that grant schools more autonomy.

In this paper we study whether the effects of a policy that expands a specific school input - instruction time - are different across two types of schools - public and no-fee charter schools - that cater to similar students but differ in the degree of autonomy. Moreover, as the the returns to additional instruction hours could depend not only on the absolute quality of time use at school, but also on its relative quality with respect to time spent outside of school, we examine whether the effects of longer schedules are heterogeneous across schoolchildren from different backgrounds.

We address these questions by exploiting the introduction of the Full School Day (FSD) in Chilean publicly subsidized (*i.e.* public and charter) primary schools. We leverage within-school variation in years of exposure to longer schedules by the end of grade 4 across several cohorts of pupils starting primary education in a given school between 2002 and 2010, and taking a standardized test at the end of grade 4. To limit the confounding effect of changes in the characteristics of pupil intake, we restrict our attention to cohorts of *incumbent* pupils, *i.e.* students who started primary education in schools that had not yet adopted the FSD and who possibly became exposed to it at some stage of their first four years of education. Furthermore, we account for potentially endogenous mobility across schools after first grade by instrumenting actual exposure with the exposure a student would have experienced, had he/she never transferred out of the school where he/she attended first grade.

[46]Bellei (2009) and Berthelon et al. (2016) find that the effect of the FSD on achievement is larger in public schools. Bellei (2009) focuses on pupils attending grade 10 in 2001 and in 2003, while Berthelon et al. (2016) examine the effect of the FSD at grade 2 in a single year (2012). In both cases, the sample of schools and the population studied are different from the ones we investigate, and the sample periods mostly do not overlap. Furthermore, to avoid that our estimates are confounded by differences in students' characteristics, we compare public schools only to charter schools that do not charge tuition fees, because we provide evidence that they serve pupils from similar backgrounds. Additionally, we show that the larger effect of the FSD found for students starting primary school in no-fee charter schools persists once we also allow longer schedules to affect schoolchildren of higher and lower socio-economic status in a different way.

We first document a positive and modest average effect of longer schedules on reading and mathematics test scores, although estimates for the latter are not statistically significant. It is important to highlight that relying on legacy enrollment cohorts for the sake of identification implies that we are able to study only the first years after the implementation of the FSD. Insofar as it takes time to adjust, short-term effects - which are interesting to study because most large-scale input expansion programs would entail some initial adaptation challenges - may be lower than long-term ones.

In order to study whether school institutions affect returns to additional instruction time, we then compare pupils who started primary education in no-fee charter schools with students who enrolled in public schools. We find that reading scores increase significantly more for the former than for the latter. We focus on these two types of establishments because they serve similar students, which attenuates concerns that the estimates reflect differences in the characteristics of students who attend different types of schools. No-fee charter schools enjoy more autonomy over the design of the curriculum and personnel decisions. Since survey evidence suggests that public and charter schools allocate additional instruction time across subjects in a similar way, we turn to analyze staff decisions and we uncover a significant difference. No-fee charter schools and public schools increased the total contract and teaching hours required to provide longer schedules in a different way: no-fee charter schools relied more on hiring additional teachers and less on increasing working hours per teacher than public schools. If public school teachers deal more often with a higher workload, which would be consistent with them reporting a lower satisfaction with the FSD, then the absolute quality of time use in public school may be lower, thus contributing to explaining the lower benefits we document for them.

Our findings are in line with the growing literature showing that charter schools are associated with improved learning outcomes (see Section 3.2). Lavy (2015) documents that the productivity of instructional time is larger in schools that have more autonomy over staff and budget decisions. We find a similar result and provide suggestive evidence that autonomy over personnel decisions seems important when expanding instruction time, because providing longer schedules requires adjusting the teaching input. In general, our results suggest that school institutions and governance matter for the effectiveness of various education policies. Further analysis on complementarities between school inputs and institutions could be a promising avenue for future research.

We also show that longer schedules do not raise significantly the reading scores of students from more advantaged backgrounds, as captured by parental education, the number of books at home and the availability of a computer and a connection to the Internet at home. On the other hand, reading scores of other pupils in the sample improve significantly. Although large in size, the difference in the benefits enjoyed by individuals from different backgrounds is not statistically significant. With this caveat in mind, our findings suggest that the relative quality of additional time spent at school is higher for pupils with fewer resources at home, thus yielding larger benefits. According to information contained in teachers' surveys, the adoption of the FSD is likely associated with a reduced frequency of homework. In light of the limited support, as shown by time-use surveys, that pupils of low socio-economic status receive outside of school, substituting autonomous study at home with supervised learning at school may be a mechanism underlying the documented heterogeneity.

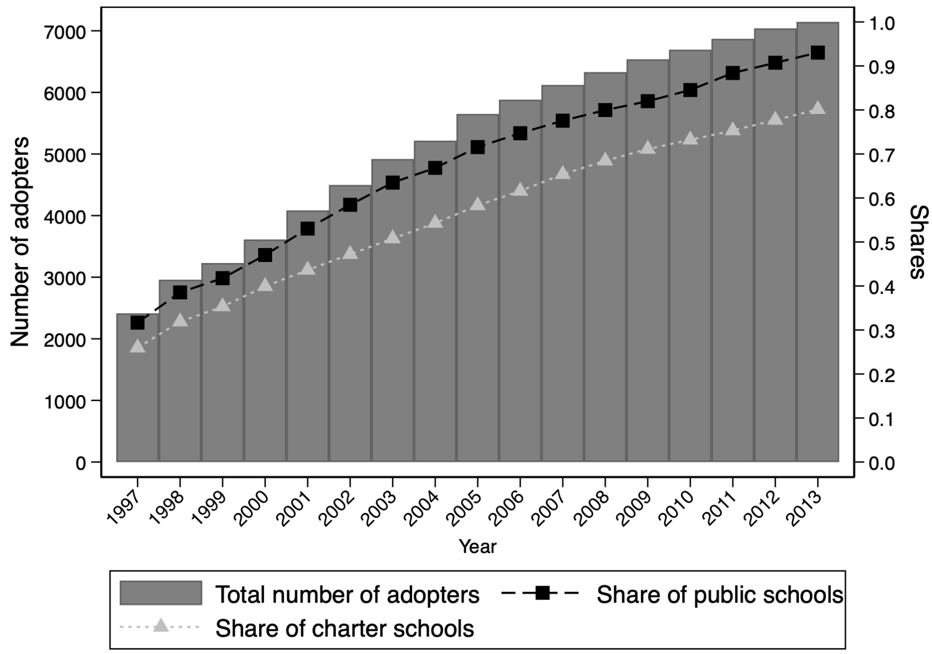
These findings are in line with the results of Lavy (2015), who presents a cross-country analysis based on 2006 PISA data.[47] If also confirmed in other settings, they would suggest that the amount of time spent at school may play a role in reducing the inequality of learning opportunities. As pupils from different backgrounds are exposed to the same school inputs for a larger part of the day, the role of household inputs - the quality of which varies greatly - may become less important. This is likely to be especially true if, as in the Chilean setting, the additional instruction time does not entail an expansion of the curriculum. Indeed, in a setting where increased weekly instruction hours serve to teach additional contents, Huebener et al. (2017) document a widening gap between high- and low-performing German pupils.

Finally, policies that extend the length of the school day may also affect non-academic outcomes. Berthelon and Kruger (2011), for instance, show that the FSD reduces the incidence of teenage pregnancies among poor girls and of youth crime; Contreras and Sepúlveda (2016) report a positive effect of the FSD on labor force participation and employment of single mothers whose youngest child is eligible to attend longer schedules. Studying these outcomes goes beyond the scope of this paper, but it is important to bear them in mind when evaluating this type of policies.

[47]When restricting the analysis to a sub-set of developing countries that include Chile, Lavy (2015) finds a stronger effect among schoolchildren from highly educated families. However, he does not provide country-specific estimates that allow to verify what is the effect in the case of Chile.

3.8 Figures

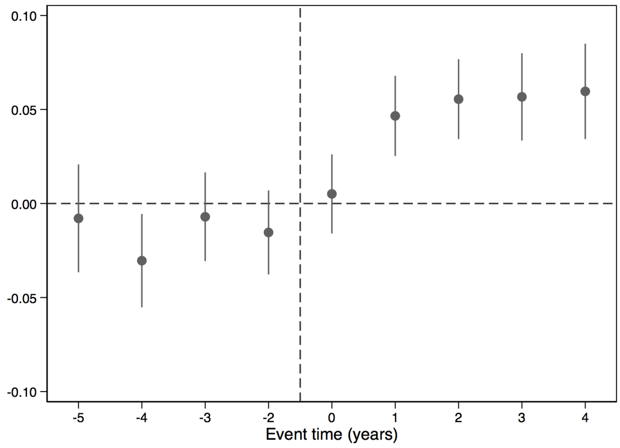
Figure 3.1: FSD adoption over the period 1997-2013



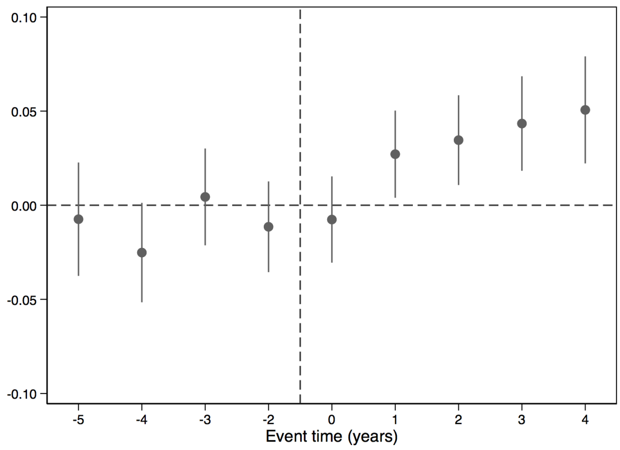
Notes: The figure illustrates the pattern of adoption of the FSD in primary publicly subsidized schools over the period 1997-2013. On the left axis it plots the number of schools that had adopted the policy by a given year; on the right axis it displays the share of public and charter schools that had implemented the FSD by a given year.

Figure 3.2: Evolution of test scores relative to 1 year before the FSD adoption

(a) Reading



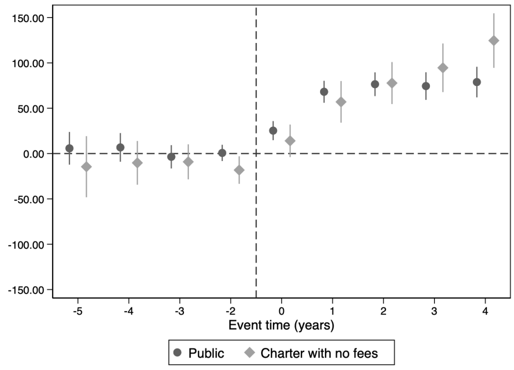
(b) Mathematics



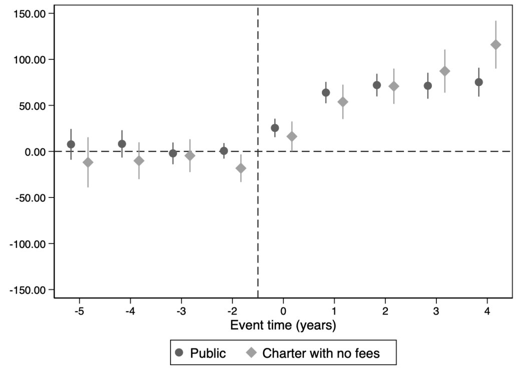
Notes: Panels (a) and (b) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification (3.4). The FSD is adopted in event-year 0 and coefficients show how different reading and mathematics test scores are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and calendar year fixed effects. Standard errors are clustered at the school level.

Figure 3.3: Evolution of contract and teaching hours relative to 1 year before the FSD adoption

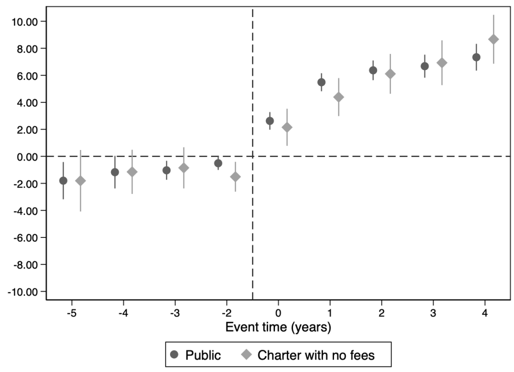
(a) Contract hours



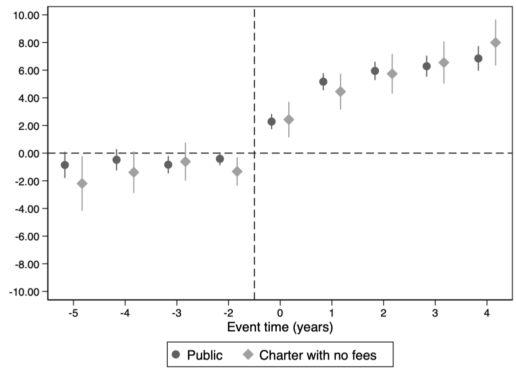
(b) Teaching hours



(c) Contract hours/n. of classes



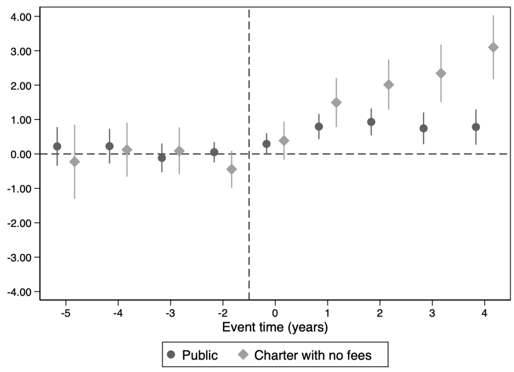
(d) Teaching hours/n. of classes



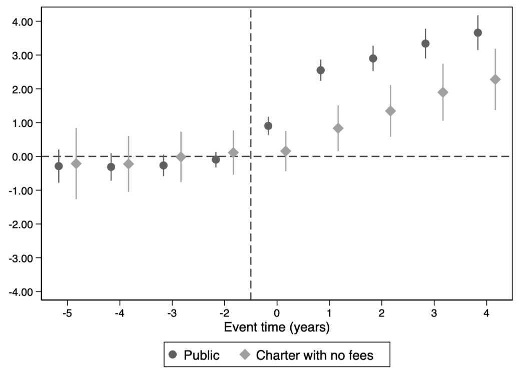
Notes: Panels (a) to (d) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (3.5). The FSD is adopted in event-year 0 and coefficients show how different contract and teaching hours are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and calendar year fixed effects. Standard errors are clustered at the school level.

Figure 3.4: Evolution of number of teachers, contract hours per teacher and teaching hours per teacher relative to 1 year before the FSD adoption

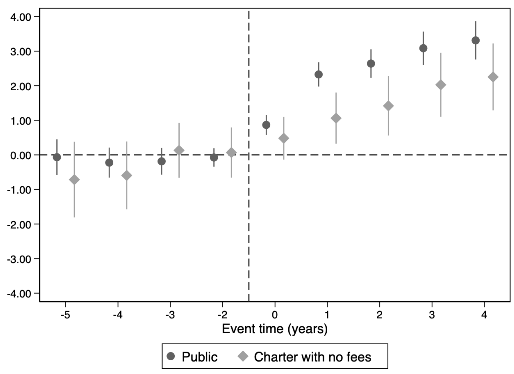
(a) Number of teachers



(b) Contract hours per teacher



(c) Teaching hours per teacher



Notes: Panels (a) to (c) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (3.5). The FSD is adopted in event-year 0 and coefficients show how different the number of teachers, contract hours per teacher and teaching hours per teacher are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and calendar year fixed effects. Standard errors are clustered at the school level.

3.9 Tables

Table 3.1: Daily schedules with and without the FSD

	FSD (1)	No FSD (2)
Minimum number of hours per day	7.08	5.00
Example of daily schedule	5 days under FSD: 08:00-15.05 4 days under FSD: 08.00 - 15.45	Morning shift: 08:00-13:00 Afternoon shift: 14:00-19:00

Notes: The table reports the minimum number of hours students spend at school every day and the daily schedule with and without the FSD in place, inclusive of time devoted to breaks. The minimum number of hours is prescribed in the law. Schools can freely choose the time at which the school day starts. The daily schedules are built assuming that the full school day and the morning shift start at 8.00, while the afternoon shift starts at 14.00.

Table 3.2: Hours of instruction per week and student voucher with and without the FSD

Subject/Grades	1st - 4th	
	FSD	No FSD
	(1)	(2)
Mathematics	6	6
Spanish	6	6
Natural and Social Sciences	6	6
Physical Education	3	3
Arts and Music	4	4
Technology	3	3
Others	2	2
School Free Choice	8	0
Total	38	30
Student Voucher (U.S.E.)	1.99	1.45

Notes: The table reports weekly subject-specific and total instruction time with and without the FSD, for grades 1 to 4. The information comes from the Decree 625 of the Ministry of Education published in 2003 (`http://bcn.cl/253tx`). It also reports the amount of the student voucher with and without the FSD, expressed in educational subsidy units (U.S.E). This information comes from the version of the DFL2/1996 of the Ministry of Education published in May, 2003 (`http://bcn.cl/1uy40`). These units underwent some modifications since the implementation of the FSD reform.

Table 3.3: Use of time under the FSD in primary schools (hours per week)

Subject	All Schools		Public Schools		Charter Schools	
	Core time	Free Choice time	Core time	Free Choice time	Core time	Free Choice time
	(1)	(2)	(3)	(4)	(5)	(6)
Spanish	5.47	2.39	5.39*	2.49	5.59*	2.24
	(0.98)	(1.64)	(0.81)	(1.59)	(1.18)	(1.71)
Mathematics	5.19	1.48	5.14	1.55	5.25	1.37
	(0.94)	(1.31)	(0.78)	(1.34)	(1.13)	(1.26)
Social Sciences	3.83	0.17	3.84	0.15	3.81	0.19
	(0.81)	(0.56)	(0.74)	(0.56)	(0.91)	(0.55)
Natural Sciences	3.89	0.49	3.91	0.47	3.85	0.51
	(0.73)	(0.94)	(0.70)	(0.93)	(0.77)	(0.96)
Foreign Languages	2.03	0.27	1.90***	0.16***	2.22***	0.43***
	(0.70)	(0.75)	(0.59)	(0.57)	(0.80)	(0.93)
Technology	2.03	0.01	2.00	0.004	2.05	0.02
	(0.53)	(0.12)	(0.52)	(0.07)	(0.54)	(0.18)
Art	3.12	0.06	3.09	0.07	3.17	0.06
	(0.81)	(0.35)	(0.77)	(0.33)	(0.86)	(0.38)
Sports	2.10	0.04	2.04**	0.028	2.19**	0.06
	(0.61)	(0.27)	(0.50)	(0.21)	(0.74)	(0.34)
Religion	1.92	0.04	1.89	0.00***	1.97	0.10***
	(0.47)	(0.28)	(0.51)	(0.00)	(0.38)	(0.43)
Number of Schools	387		229		158	

Notes: The table reports hours per week allocated to different subjects in fifth grade for a representative sample of urban schools that had adopted the FSD by 2005 and were surveyed by the Studies Directorate of the Sociology Faculty at the Catholic University of Chile (*DESUC*). "Core" time excludes "Free Choice" time. *, **, *** indicate that the number of hours allocated to a given subject is significantly different between public and charter schools at the 10, 5 and 1 per cent level, respectively. Standard deviations are in parentheses.

Table 3.4: Summary statistics

	All	Public	Charter	
			No Tuition Fees	Tuition Fees
	(1)	(2)	(3)	(4)
Students demographics				
Female	0.49	0.50	0.49	0.49
Age at school entry	6.60	6.60	6.60	6.59
Parental education				
Less than university	0.87	0.92	0.91	0.78
Resources at home				
Books (1 = At most 50)	0.85	0.89	0.88	0.79
Computer (1=yes)	0.55	0.42	0.50	0.72
Internet (1 = yes)	0.31	0.20	0.27	0.46
Schools Characteristics				
First grade average class size	34.68	33.69	35.37	35.53
First grade enrollment	82.33	83.09	72.39	86.56
Academic performance				
Fourth grade reading test score	-0.04	-0.17	-0.12	0.18
Fourth grade Mathematics test score	-0.04	-0.17	-0.15	0.19
First-grade attendance rate	94.23	93.88	94.83	94.33
End of first-grade status (1=repeat)	0.03	0.03	0.04	0.02
N. of students	604532	270417	114074	218495

Notes: The table reports summary statistics for the sample of fourth graders who started primary school between 2002 and 2010 in publicly subsidized schools that had not yet adopted the FSD. Parental education refers to the highest educational attainment among the mother and the father; in case the information is missing for one parent, it refers to the education level of the other parent. All figures are expressed as fractions, except from averages referring to the age of pupils, class size, enrollment, test scores and the attendance rate. Test scores are standardized by year and subject (including also pupils who are not in the master sample) to have mean equal to 0 and standard deviation equal to 1. The number of observations in columns (2) to (4) does not sum to the number of observations in column (1) because for 2.39 per cent of charter schools we could not find information about the tuition fees.

Table 3.5: Effect of the FSD on test scores

	Linear specification				Non parametric specification	
	FE1	FE2	FE-IV1	FE-IV2	FE-IV1	FE-IV2
	(1)	(2)	(3)	(4)	(5)	(6)
				A. Reading		
Years under FSD	0.002	0.011***	0.024***	0.024***		
	(0.003)	(0.003)	(0.006)	(0.006)		
Years under FSD = 1					0.023	0.022
					(0.015)	(0.015)
Years under FSD = 2					0.029**	0.030**
					(0.015)	(0.014)
Years under FSD = 3					0.116***	0.114***
					(0.021)	(0.021)
First stage coefficient			0.720***	0.720***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			23614.36	24416.41	5136.66	5265.46
N. of students	596108	596108	596108	596108	596108	596108
				B. Mathematics		
Years under FSD	-0.007**	0.005	0.007	0.008		
	(0.003)	(0.003)	(0.007)	(0.007)		
Years under FSD = 1					-0.014	-0.015
					(0.016)	(0.016)
Years under FSD = 2					-0.003	-0.002
					(0.017)	(0.016)
Years under FSD = 3					0.057**	0.058**
					(0.023)	(0.023)
First stage coefficient			0.719***	0.720***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			23460.87	24294.13	5140.07	5278.70
N. of students	596281	596281	596281	596281	596281	596281
Student-level controls	No	Yes	No	Yes	No	Yes
School-level controls	No	Yes	No	Yes	No	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports the effect of the FSD on reading and mathematics test scores. Estimates in columns (1), (3) and (5) are based on a parsimonious specification that only includes as controls school fixed effects and year fixed effects. Estimates in other columns are based on a richer specification that features an additional set of controls. Specifically, student-level controls include: gender, age at school entry, as well as the attendance rate and the status (pass or repeat) at the end of grade 1. School-level controls include averages of the students' characteristics at the school level, as well as enrollment and average class size in first grade. The effect of the FSD is assumed to be linear in exposure in columns (1) to (4), whereas it is allowed to vary in a fully non-parametric way in columns (5) and (6). In specifications FE-IV1 and FE-IV2 the treatment (i.e. actual years of exposure to the FSD by the end of grade 4) is instrumented with the exposure a student would experience had he/she never transferred out of his/her first-grade school. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 3.6: Heterogeneous effects of the FSD on test scores by students' socio-economic background

	Reading			Mathematics		
	$D = \mathbb{1}$(No Uni.)	$D = \mathbb{1}$(Books $\leq$ 50)	$D = \mathbb{1}$(No ICT)	$D = \mathbb{1}$(No Uni.)	$D = \mathbb{1}$(Books $\leq$ 50)	$D = \mathbb{1}$(No ICT)
	(1)	(2)	(3)	(4)	(5)	(6)
Years under FSD	0.007	0.012	0.012	-0.003	0.003	0.004
	(0.012)	(0.011)	(0.009)	(0.012)	(0.012)	(0.010)
Years under FSD $\times$ D	0.015	0.010	0.012	0.011	0.005	0.004
	(0.013)	(0.011)	(0.010)	(0.012)	(0.011)	(0.011)
Kleibergen-Paap rk Wald F statistic	2781.84	11502.69	5644.29	2797.02	11528.94	5643.12
N. of students	532970	529879	517249	534473	531369	518676
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows the effect of the FSD on reading and mathematics test scores by different measures of students' socio-economic background. Every column shows an enriched version of specification (3.2) where the treatment and all controls listed in the notes to Table 3.5, including school and year fixed effects, are also interacted with a dummy D, capturing a relevant dimension of heterogeneity. In columns (1) and (4) D takes value 1 if no parent in the household has some university education, and 0 otherwise. In columns (2) and (5) D takes value 1 if there are at most 50 books at home, and 0 otherwise. In columns (3) and (6) D takes value 1 if there is not a computer or an Internet connection at home, and 0 otherwise. Actual years of exposure to the FSD are instrumented with years of exposure a student would accumulate had he/she never transferred from the school where he/she attended first grade. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 3.7: Heterogeneous effects of the FSD on test scores by school type

	Reading				Mathematics			
		$D = \mathbb{1}$(No Uni.)	$D= \mathbb{1}$(Books ≤ 50)	D=$\mathbb{1}$(No ICT)		$D = \mathbb{1}$(No Uni.)	D=$\mathbb{1}$(Books ≤ 50)	D=$\mathbb{1}$(No ICT)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Years under FSD	0.019**	-0.001	0.023	0.011	0.012	-0.007	0.008	0.003
	(0.009)	(0.021)	(0.017)	(0.015)	(0.009)	(0.021)	(0.017)	(0.016)
Years under FSD × no-fee charter	0.042**	0.047**	0.046**	0.045**	0.017	0.017	0.015	0.018
	(0.019)	(0.019)	(0.019)	(0.020)	(0.021)	(0.021)	(0.021)	(0.020)
Years under FSD × SES char. D		0.022	-0.003	0.009		0.024	0.010	0.015
		(0.021)	(0.017)	(0.015)		(0.020)	(0.017)	(0.015)
Kleibergen-Paap rk Wald F statistic	828.30	605.18	609.12	632.25	839.28	602.71	606.90	630.61
N. of students	377856	319187	319234	319254	377719	320061	320108	320129
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table shows the effect of the FSD on reading and mathematics test scores by the type of school (public or charter without tuition fees) that the student attended in grade 1. Estimates in columns (1) and (5) are based on an enriched version of specification (3.2), where the treatment and all controls listed in the notes to Table 3.5, including year fixed effects, are also interacted with a dummy taking value 1 if the school is a charter establishment without tuition fees, and 0 if it is public. In columns (2) and (6) the treatment and all controls are further interacted with a dummy that takes value 1 if none of the parents has some university education, and 0 otherwise. In columns (3) and (7) the treatment and all controls are further interacted with a dummy that takes value 1 if there are at most 50 books at home, and 0 otherwise. In columns (4) and (8) the treatment and all controls are further interacted with a dummy that takes value 1 if either the computer or a connection to the Internet are not available at home, and 0 otherwise. Actual years of exposure to the FSD are instrumented with years of exposure a student would accumulate had he/she never transferred from the school where he/she attended first grade. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Appendices

3.A Public and Charter Schools

Public and charter schools are subject to different regulations. This translates into charter schools enjoying more autonomy and flexibility over curriculum, budgetary and personnel decisions.

Public schools are either managed by the Municipal Department of Education (*DAEM*) or by private non-profit education corporations. While the director of the *DAEM* is a teacher, corporations are led by a board of directors who do not need to be teachers and whose president is the major of the municipality. Under both management schemes decisions related to the allocation of resources and to hiring/firing school staff are taken at the municipality level and school principals are not necessarily involved. Charter schools are instead private organizations and all relevant decisions are taken by the school authorities.

The working conditions of the employees of public schools are regulated by the *Estatuto de los Profesionales de la Educación*. The relevant regulation for charter schools is the *Código del Trabajo*, the labor code that applies to all firms in Chile. Appointments of public school teachers are decided by a commission that is formed by the Major, the director of the *DAEM* or the education corporations, as well as one randomly selected teacher from the schools in the municipality. Priority is given to spouses of teachers already working in the municipality. The salary of public school teachers is fixed according to a national scale that takes into account experience, training, specific difficult situations (such as teaching in rural, remote or deprived areas) and responsibilities. Firing is subject to many restrictions. It is possible only if one of the following conditions are met: (i) school enrollment decreases; (ii) the national curriculum undergoes changes that justify the decision; (iii) schools' merges; (iv) protracted poor performance (see below). Teachers having tenured positions enjoy a greater job security.[48] In any case, firings have to be justified in the Annual Plan of Educational Development that needs the approval of the Provincial Office of the Ministry of Education. Charter schools are instead free to set their own recruitment and dismissal criteria. Wages and the other working

[48]The *Estatuto de los Profesionales de la Educación* contemplates two type of contracts, *titular* and *contratado*. The first type of contract affords a greater job security, as it offers a tenured position.

conditions are subject to the same regulations that apply to private firms.

There are also differences in the evaluation of teachers. The *Estatuto de los Profesionales de la Educación* originally set some criteria for assessing teachers performance, but they were never fully implemented. In 2003 a new evaluation system was agreed. Nevertheless it is quite lax and in practice very few teachers receive poor evaluations. In principle, teachers could be fired if they fare unsatisfactorily in two or three consecutive evaluations. School principals are not accountable based on the school performance and they can be fired only in case of a grave fault, while poor evaluations can result in assigning them to smaller schools. Charter schools can instead set their own evaluation systems and the consequences in case of poor performance.

3.B Robustness Checks

As a first robustness check, we show that specification (3.2) delivers similar estimates if we also control for a set of characteristics of teachers in the school s where the student started first grade in year t. Specifically, we include as controls the share of female teachers, the share of teachers with an education degree and teachers' average age.[49] Appendix Table 3.B.1 shows that results are virtually unchanged.

A possible concern not addressed by specification (3.2) is that other events may happen in a school around the time of FSD adoption and affect learning outcomes in the following years. Our estimates would then also capture the effects of other changes to the school environment.

The first potential confounder to check is infrastructure investment, as some schools had to expand their infrastructure prior to switching to a single-shift scheme. Funds disbursed for this purpose covered costs related to replicating the existing infrastructure on a larger scale, not to improving it. Nonetheless, to address this issue, we replicate our analysis on the sample of pupils who started first grade in schools that did not receive public funds for expanding infrastructure. These establishments are unlikely to have made substantial changes to their facilities prior to lengthening the school day. Columns (1) and (2) of Appendix Table 3.B.2 report estimates that are in a similar range as those coming from the full sample of schools. An additional year of exposure to the FSD raises reading test scores by 0.020σ. The effect on mathematics test scores is virtually 0. According to this exercise, infrastructure investment does not appear to be an important alternative driver of our estimates.

In 2008 Chile introduced a Preferential School Subsidy scheme (*Subvención Escolar Preferencial*, or PSS henceforth) which grants schools an additional subsidy for each disadvantaged student they cater to.[50] To check whether our estimates are also capturing the roll-out of the subsidy, we implement two exercises. First, we enrich specification (3.2) with controls for the individual exposure to the PSS scheme (*i.e.* the number of grades during which the student received the subsidy) by grade 4 and the average share of pupils benefiting from the PSS scheme in the schools attended by a student in grades 1 to 4. Second, we estimate specification (3.2) on the sub-sample of cohorts never exposed to the PSS (*i.e.* those starting primary education before 2005). In both cases, coefficients are similar to those coming from the main specification and, if anything, in the case of reading they are slightly larger

[49] These controls are not included in the baseline specification because they are not available for the year 2002 and are missing for some schools in other years. In this regression specification, we assume that the teaching staff in 2002 is the same as that observed in 2003, so as not to drop one year of observations.

[50] The receipt of the subsidy is conditional upon schools developing a pedagogical plan that outlines how additional funds are used to improve learning outcomes and upon allowing for an external evaluation of the results achieved. See Santiago et al. (2013) for more information.

(Appendix Table 3.B.2, columns 3 and 5).

Table 3.B.1: Effect of the FSD on test scores including teacher controls

	Linear specification				Non parametric specification	
	FE2	FE2	FE-IV2	FE-IV2	FE-IV2	FE-IV2
	(1)	(2)	(3)	(4)	(5)	(6)
			A. Reading			
Years under FSD	0.011***	0.011***	0.024***	0.024***		
	(0.003)	(0.003)	(0.006)	(0.006)		
Years under FSD = 1					0.022	0.023
					(0.015)	(0.015)
Years under FSD = 2					0.030**	0.032**
					(0.014)	(0.014)
Years under FSD = 3					0.114***	0.115***
					(0.021)	(0.021)
First stage coefficient			0.720***	0.723***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			24416.41	24689.20	5265.46	5244.84
N. of students	596108	578112	596108	578112	596108	578112
			B. Mathematics			
Years under FSD	0.005	0.004	0.008	0.009		
	(0.003)	(0.003)	(0.007)	(0.007)		
Years under FSD = 1					-0.015	-0.012
					(0.016)	(0.016)
Years under FSD = 2					-0.002	0.001
					(0.016)	(0.016)
Years under FSD = 3					0.058**	0.060***
					(0.023)	(0.023)
First stage coefficient			0.720***	0.723***		
			(0.005)	(0.005)		
Kleibergen-Paap rk Wald F statistic			24294.13	24506.65	5278.70	5253.62
N. of students	596281	578281	596281	578281	596281	578281
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Teacher-level controls	No	Yes	No	Yes	No	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports the effect of the FSD on reading and mathematics test scores. Estimates in columns (1), (3) and (5) are based on the specification with baseline controls. Student-level controls include: gender, age at school entry, as well as the attendance rate and the status (pass or repeat) at the end of grade 1. School-level controls include averages of the students' characteristics at the school level, as well as enrollment and average class size. Estimates in columns (2), (4) and (6) include also controls referring to teachers' characteristics when the students attend grade 1. Specifically, they are the share of female teachers, teachers' average age and the share of teachers with an education degree. The treatment in specifications FE2 is actual years of exposure to the FSD by the end of grade 4, while in specifications FE-IV2 is instrumented with the exposure a student would experience had he/she never transferred out of his/her first grade school. Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 3.B.2: Effect of the FSD on test scores - robustness checks

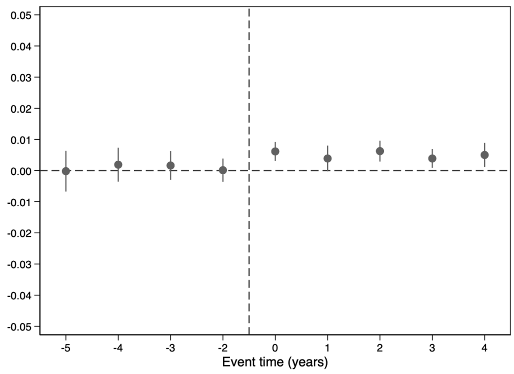
	No infrastructure funds		PSS			
	Reading	Mathematics	Reading	Mathematics	Reading	Mathematics
	(1)	(2)	(3)	(4)	(5)	(6)
Years under FSD	0.020**	-0.001	0.027***	0.011	0.026**	0.002
	(0.010)	(0.011)	(0.006)	(0.007)	(0.012)	(0.013)
Number of students	379449	379691	596020	596190	291057	291085
Kleibergen-Paap rk Wald F statistic	9202.99	9259.78	21528.75	21474.33	11287.84	11198.08
Student-level controls	Yes	Yes	Yes	Yes	Yes	Yes
School-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table presents the results from a set of specifications that check the robustness of the main estimates of the effects of the FSD on test scores. All specifications include school and year fixed effects, and actual exposure to the FSD is instrumented with the exposure a student would accumulate had he/she never transferred from the school where he/she attended first grade. In columns (1) and (2), specification (3.2) is estimated on the sub-sample of pupils in the master sample who start first grade in schools that did not receive public funds for expanding their infrastructure. In columns (3) and (4) specification (3.2) is enriched with two additional controls, on top of those listed in the notes to Table 3.5: individual exposure to the Preferential Subsidy Scheme (PSS) policy by grade 4 and the average share of pupils benefiting from the PSS in the schools attended by the student in grades 1 to 4. In columns (5) and (6) , specification (3.2) is estimated on the sub-sample of cohorts never exposed to the Preferential Subsidy Scheme (i.e. cohorts starting primary education between 2002 and 2004). Standard errors are clustered at the school level and are reported in parenthesis.
*, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

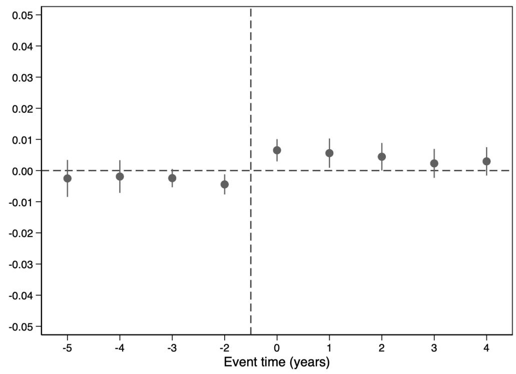
3.C Additional Figures and Tables

Figure 3.C.1: Evolution of the number of subjects relative to 1 year before the FSD adoption

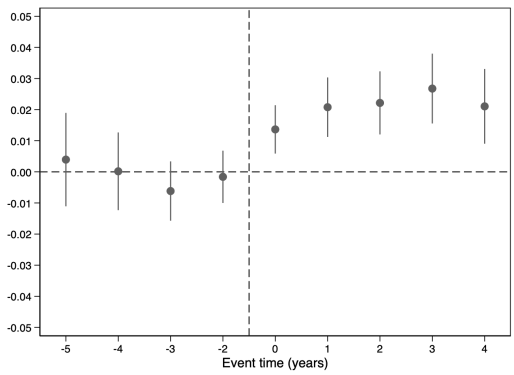
(a) Spanish



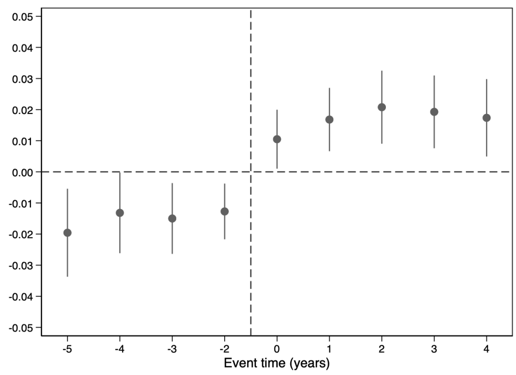
(b) Mathematics



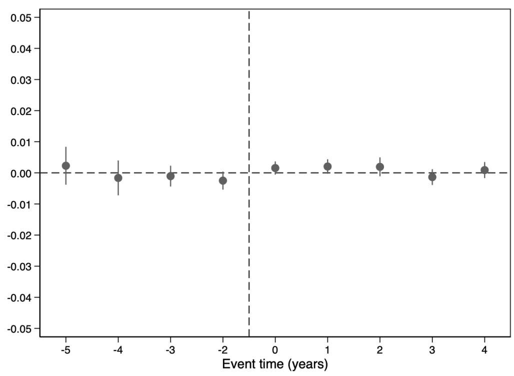
(c) Foreign languages



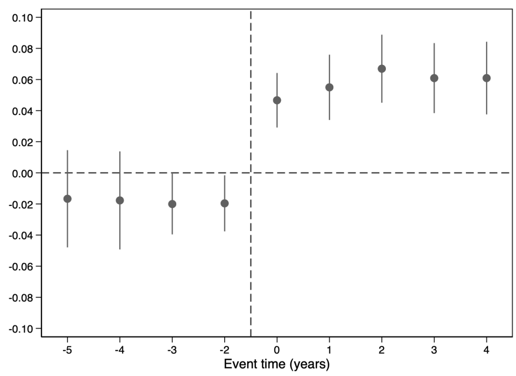
(d) Tutorials and workshops



(e) Sport



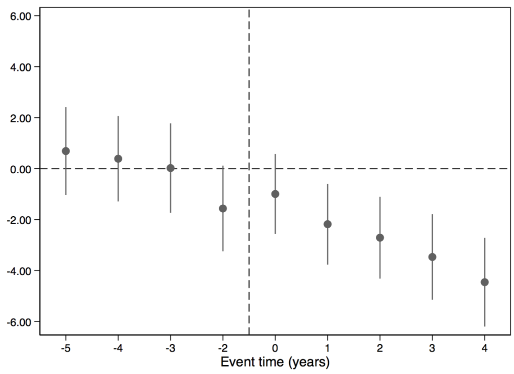
(f) Total n. of subjects



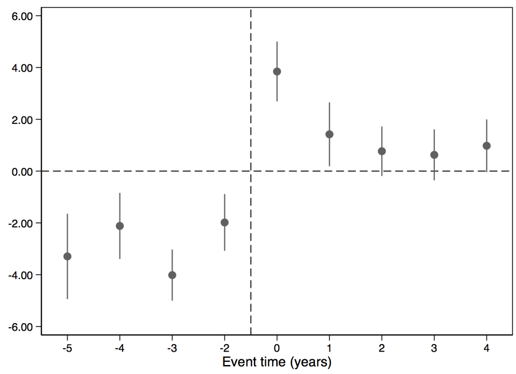
Notes: Panels (a) to (f) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (3.1). The FSD is adopted in event-year 0 and coefficients show how different the number of subjects taught is in event-year p relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school, grade and calendar year fixed effects, as well as their interactions. Standard errors are clustered at the school-grade level.

Figure 3.C.2: Evolution of transfers relative to 1 year before the FSD adoption

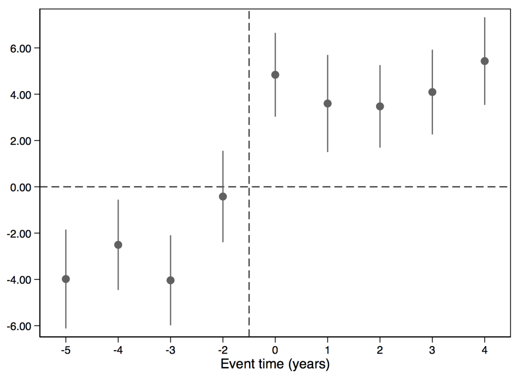
(a) Transfers out of the school



(b) Transfers into the school

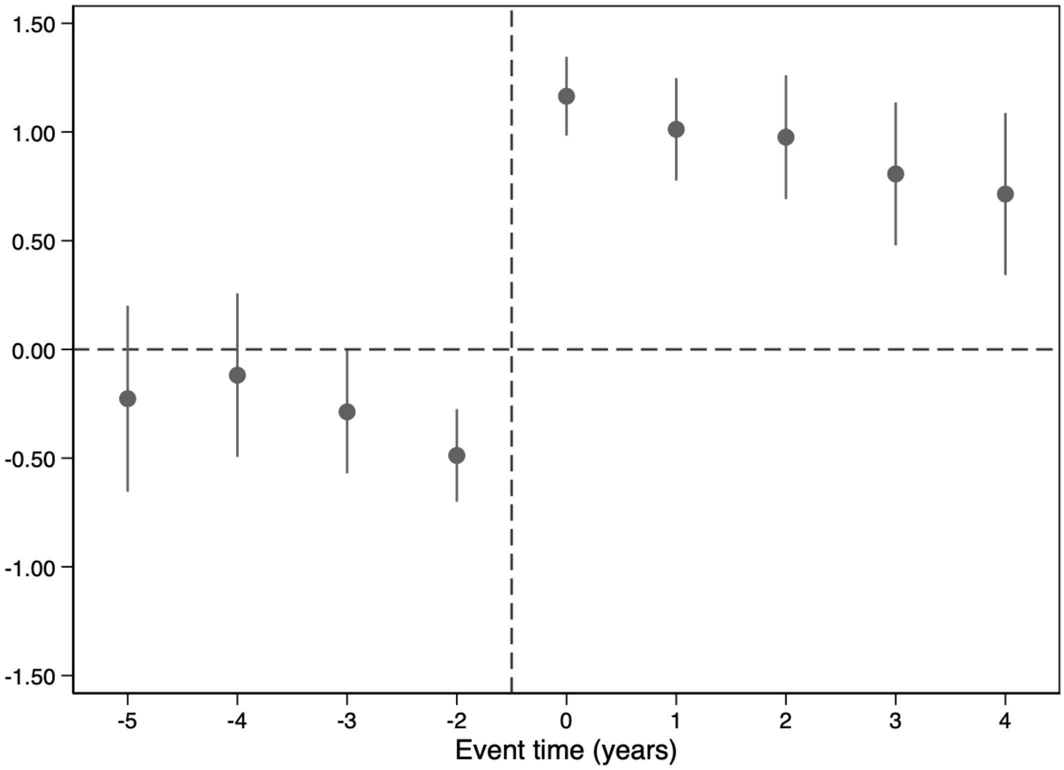


(c) Net transfers (in-out)



Notes: Panels (a) to (c) plot coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (3.5). The FSD is adopted in event-year 0 and coefficients show how different the number of transfers in grades 1 to 4 is in event-year p relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and calendar year fixed effects. Standard errors are clustered at the school level.

Figure 3.C.3: Evolution of class size relative to 1 year before the FSD adoption



Notes: The figure plots coefficients, alongside 95 per cent confidence intervals, from the event study specification outlined in (3.5). The FSD is adopted in event-year 0 and coefficients show how different the average class size in grades 1 to 4 is in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and calendar year fixed effects. Standard errors are clustered at the school level.

Table 3.C.1: Characteristics of students in the master sample who do and do not transfer

	First grade Academic Performance		FSD Exposure	
	Attendance (1)	Repetition (2)	Real (3)	Potential (4)
Do not transfer between grades 1 and 4	94.64	0.02	0.52	0.51
Transfer between grades 1 and 4	93.44	0.05	1.35	0.42

Notes: Columns (1) and (2) show the average attendance rate in grade 1 for students in the master sample and the fraction of them who repeat grade 1, distinguishing pupils who never transfer between grades 1 and 4 from those who transfer. Columns (3) and (4) display their average actual exposure to the FSD by grade 4 as well their average "potential" exposure, i.e. the years of exposure a student would experience had he/she never transferred out of his/her first-grade school.

Table 3.C.2: Support received by students outside of school

	Help with homework from household head and head' spouse		Tutoring	
	1 = Yes	Hours	1 = Yes	Hours
	(1)	(2)	(3)	(4)
No university	0.328	0.462	0.050	0.123
University	0.484	0.717	0.119	0.280
Observations	6205	6205	2671	2671

Notes: The table shows the amount of support that students receive outside of school, depending on whether they live in a household where one among the household head and the head' spouse has some university education (row "University") or not (row "No university"). The units of observation are the households heads and their spouses (in households where these is at least one child aged 5-18) when the question is whether they provide help with homework. The units of observations are pupils aged 12-18 (younger children are not interviewed) when the question is whether they receive tutoring outside of school. Information is drawn from the 2015 Chilean Time-Use Survey (*Encuesta nacional sobre uso del tiempo*).

Table 3.C.3: Frequency of mathematics homework

	A. All schools		
	No FSD	FSD	
	(1)	(2)	(3)
Every class	20.04%	11.80%	
Almost every class	50.39%	39.35%	
Some classes	28.42%	46.33%	
Never	1.15%	2.51%	
N. of Teachers	3294	18494	
	B. Schools that had not adopted the FSD by 2011		
	2011	2013	
		No FSD	FSD
Every class	22.51%	18.81%	7.95%
Almost every class	51.88%	52.10%	31.82%
Some classes	24.50%	28.34%	59.09%
Never	1.11%	0.74%	1.14%
N. of Teachers	902	808	88

Notes: The table reports information about the frequency of mathematics homework, drawn from the 2011, 2012 and 2013 waves of the teacher surveys administered alongside the SIMCE test. Panel A compares the frequency of homework in schools with and without the FSD. Panel B focuses on schools that had not adopted the FSD by 2011 and compare homework frequency in 2011 and 2013. In 2013, schools are divided according to whether they switched to longer schedules by that year (column 3) or not (column 2).

Table 3.C.4: Differences in school autonomy between public and charter Schools

	Public schools (1)	Charter schools (2)
Textbook use	98	100
Courses content	30	63
Courses offer	70	97
Formulate budget	18	96
Allocate budget	52	97
Hire teachers	28	98
Fire teachers	11	97
Set starting salaries	2	88
Increase salaries	2	91
Observations	62	85

Notes: The table reports the percentage of schools in which the principal, the teachers or the governing body have a considerable responsibility over the listed tasks. Information comes from the 2006 and 2009 school surveys administered alongside PISA tests. The sample consists of all public or charter schools in the Chilean PISA sample that also offer primary education.

Table 3.C.5: Evolution of teacher related inputs relative to 1 year before the FSD adoption

	Contract HH.	Teaching HH.	$\frac{\text{Contract HH.}}{\text{N. of Classes}}$	$\frac{\text{Teaching HH.}}{\text{N. of Classes}}$	N. of Teachers	$\frac{\text{Contract HH.}}{\text{N. of Teachers}}$	$\frac{\text{Teaching HH.}}{\text{N. of Classes}}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Event-year -5	6.507	8.355	-1.838***	-0.861*	0.239	-0.285	-0.058
	(9.269)	(8.588)	(0.710)	(0.489)	(0.289)	(0.253)	(0.266)
Event-year -4	6.712	8.316	-1.198*	-0.481	0.222	-0.309	-0.209
	(8.060)	(7.596)	(0.622)	(0.396)	(0.257)	(0.209)	(0.223)
Event-year -3	-3.556	-1.857	-1.058***	-0.850***	-0.115	-0.262	-0.187
	(6.566)	(6.063)	(0.360)	(0.325)	(0.214)	(0.165)	(0.198)
Event-year -2	1.630	1.301	-0.539**	-0.451*	0.073	-0.085	-0.072
	(4.568)	(4.302)	(0.252)	(0.242)	(0.150)	(0.115)	(0.137)
Event-year 0	25.898***	26.016***	2.570***	2.222***	0.297*	0.921***	0.871***
	(5.359)	(5.119)	(0.334)	(0.280)	(0.160)	(0.139)	(0.149)
Event-year 1	69.697***	65.191***	5.473***	5.148***	0.832***	2.567***	2.325***
	(6.157)	(5.886)	(0.342)	(0.318)	(0.186)	(0.161)	(0.180)
Event-year 2	77.471***	72.651***	6.323***	5.891***	0.951***	2.912***	2.636***
	(6.671)	(6.254)	(0.375)	(0.340)	(0.199)	(0.192)	(0.212)
Event-year 3	75.693***	72.580***	6.638***	6.289***	0.777***	3.339***	3.102***
	(7.754)	(7.170)	(0.437)	(0.392)	(0.235)	(0.226)	(0.245)
Event-year 4	80.204***	76.055***	7.345***	6.837***	0.815***	3.674***	3.303***
	(8.578)	(7.968)	(0.507)	(0.456)	(0.261)	(0.264)	(0.284)
Event-year -5 · no-fee charter	-21.642	-20.688	-0.014	-1.378	-0.490	0.060	-0.640
	(19.645)	(16.428)	(1.369)	(1.127)	(0.624)	(0.598)	(0.621)
Event-year -4 · no-fee charter	-17.450	-18.870	0.019	-0.933	-0.118	0.066	-0.384
	(14.774)	(12.779)	(1.046)	(0.865)	(0.478)	(0.473)	(0.550)
Event-year -3 · no-fee charter	-5.987	-3.150	0.213	0.244	0.184	0.229	0.320
	(11.892)	(11.005)	(0.861)	(0.777)	(0.411)	(0.415)	(0.450)
Event-year -2 · no-fee charter	-20.465**	-20.247**	-0.978	-0.876	-0.543*	0.167	0.119
	(9.075)	(8.880)	(0.624)	(0.580)	(0.316)	(0.354)	(0.397)
Event-year 0 · no-fee charter	-11.967	-9.823	-0.261	0.359	0.077	-0.754**	-0.370
	(10.717)	(9.832)	(0.778)	(0.714)	(0.328)	(0.338)	(0.353)
Event-year 1 · no-fee charter	-12.675	-11.155	-1.068	-0.645	0.654	-1.804***	-1.331***
	(13.391)	(11.319)	(0.798)	(0.736)	(0.415)	(0.379)	(0.413)
Event-year 2 · no-fee charter	0.834	-1.541	-0.223	-0.166	1.058**	-1.553***	-1.221**
	(13.714)	(11.724)	(0.826)	(0.791)	(0.426)	(0.435)	(0.488)
Event-year 3 · no-fee charter	19.009	14.212	0.236	0.174	1.549***	-1.475***	-1.147**
	(15.902)	(14.086)	(0.961)	(0.877)	(0.493)	(0.489)	(0.536)
Event-year 4 · no-fee charter	45.435**	40.769***	1.375	1.244	2.308***	-1.464***	-1.122**
	(17.756)	(15.615)	(1.065)	(0.973)	(0.547)	(0.536)	(0.571)
School fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Calendar year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N. of school-years	18765	18765	18538	18538	18765	18765	18765

Notes: The table reports coefficients from a richer version of the event study specification outlined in (3.5) where calendar year fixed effects and event years are also interacted with a dummy D_s taking value 1 if the school is a no-fee charter school, and 0 otherwise. The FSD is adopted in event-year 0 and coefficients show how different total contract hours, teaching hours, contract hours per class, teaching hour per class, total number of teachers, contract hours per teacher and teaching hours per teacher are in event-year ρ relative to event-year -1, which is taken as the reference year. The sample consists of all schools where students in the master sample enrolled in first grade. All specifications include school and calendar year fixed effects. Standard errors are clustered at the school level. *, ** and *** denote significance at the 10, 5 and 1 per cent level, respectively.

Table 3.C.6: Teachers opinion about the FSD

	Public Schools (1)	Charter Schools (2)
Good or very good	44.99%	54.44%
Not bad, not good	32.99%	29.89%
Bad or very bad	22.02%	15.67%

Notes: The table reports the opinion of teachers about the FSD, dividing them according to the school (public or charter) in which they teach. Information is drawn from the 2005 wave of the *Encuesta Longitudinal Docente* implemented by the *Centro de Microdatos* of the *Universidad de Chile*.

References

Abdulkadiroğlu, Atila, Joshua D Angrist, Peter D Hull, and Parag A Pathak, "Charters without lotteries: testing takeovers in New Orleans and Boston," *The American Economic Review*, 2016, *106* (7), 1878–1920.

_ , _ , Susan M Dynarski, Thomas J Kane, and Parag A Pathak, "Accountability and flexibility in public schools: evidence from Boston's Charters and Pilots," *The Quarterly Journal of Economics*, 2011, *126* (2), 699–748.

Acemoglu, Daron and Pascual Restrepo, "Demographics and automation," *Working paper*, 2019.

Agüero, Jorge M and Trinidad Beleche, "Test-Mex: estimating the effects of school year length on student performance in Mexico," *Journal of Development Economics*, 2013, *103*, 353–361.

Almeida, Rita, Antonio Bresolin, Bruna Pugialli Da Silva Borges, Karen Mendes, and Naercio A Menezes-Filho, "Assessing the impacts of Mais Educacao on educational outcomes: evidence between 2007 and 2011," *World Bank Policy Research Working Paper 7644*, 2016.

Angrist, Joshua D, Parag A Pathak, and Christopher R Walters, "Explaining charter school effectiveness," *American Economic Journal: Applied Economics*, 2013, *5* (4), 1–27.

Atalay, Kadir and Garry F Barrett, "The impact of age pension eligibility age on retirement and program dependence: evidence from an Australian experiment," *Review of Economics and Statistics*, 2015, *97* (1), 71–87.

Attanasio, Orazio P and Agar Brugiavini, "Social security and households' saving," *the Quarterly Journal of economics*, 2003, *118* (3), 1075–1119.

Aucejo, Esteban M and Teresa Foy Romano, "Assessing the effect of school days and absences on test score performance," *Economics of Education Review*, 2016, *55*, 70–87.

Auerbach, Alan J and Laurence J Kotlikoff, "The efficiency gains from Social Security benefit-tax linkage," *NBER Working Paper no. 1645*, 1985.

Baker, George, Michael Gibbs, and Bengt Holmstrom, "The internal economics of the firm: Evidence from personnel data," *The Quarterly Journal of Economics*, 1994, *109* (4), 881–919.

Baltrunaite, Audinga, Cristina Giorgiantonio, Sauro Mocetti, and Tommaso Orlando, "Discretion and supplier selection in public procurement," *Banca d'Italia, Temi di Discussione*, 2018, *No. 1178.*

Banca d'Italia, "Relazione Annuale sul 1993," Technical Report 1993.

Banks, James, Rowena Crawford, and Gemma Tetlow, "Annuity choices and income drawdown: evidence from the decumulation phase of defined contribution pensions in England," *Journal of pension economics & finance*, 2015, *14* (4), 412–438.

Battistin, Erich and Elena Claudia Meroni, "Should we increase instruction time in low achieving schools? Evidence from Southern Italy," *Economics of Education Review*, Dec 2016, *55*, 39–56.

Baude, Patrick L, Marcus Casey, Eric A Hanushek, Gregory R Phelan, and Steven G Rivkin, "The evolution of charter school quality," *Economica*, 2014.

Behaghel, Luc and David M Blau, "Framing social security reform: Behavioral responses to changes in the full retirement age," *American Economic Journal: Economic Policy*, 2012, *4* (4), 41–67.

Bellei, Cristián, "Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile," *Economics of Education Review*, 2009, *28* (5), 629–640.

Berthelon, Matias, Diana I Kruger, and Veronica Vienne, "Longer school schedules and early reading skills: effects from a full-day school reform in Chile," *IZA DP n. 10282*, 2016.

Berthelon, Matias E and Diana I Kruger, "Risky behavior among youth: incapacitation effects of school on adolescent motherhood and crime in Chile," *Journal of public economics*, 2011, *95* (1), 41–53.

Bertoni, Marco and Giorgio Brunello, "Does delayed retirement affect youth employment? Evidence from Italian local labour markets," *Working paper*, 2017.

_ , _ , and Gianluca Mazzarella, "Does postponing minimum retirement age improve healthy behaviors before retirement? Evidence from middle-aged Italian workers," *Journal of health economics*, 2018, *58*, 215–227.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How much should we trust differences-in-differences estimates?," *The Quarterly journal of economics*, 2004, *119* (1), 249–275.

Betts, Julian R and Eric Johnson, "A test of diminishing returns to school spending," *mimeographed, University of California San Diego*, 1998.

Bianchi, Nicola, Giulia Bovini, Jin Li, Matteo Paradisi, and Michael Powell, "Career spillovers in internal labour market," *Working paper*, 2019.

Boeri, Tito, Pietro Garibaldi, and Espen Moen, "Closing the retirement door and the lump of labor," *Working paper*, 2017.

Borghans, Lex, Anne C Gielen, and Erzo FP Luttmer, "Social support substitution and the earnings rebound: Evidence from a regression discontinuity in disability insurance reform," *American economic Journal: economic policy*, 2014, *6* (4), 34–70.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental shift-share research design," *Working paper*, 2018.

Bottazzi, Renata, Tullio Jappelli, and Mario Padula, "Retirement expectations, pension reforms, and their impact on private wealth accumulation," *Journal of Public Economics*, 2006, *90* (12), 2187–2212.

_ , _ , and _ , "The portfolio effect of pension reforms: evidence from Italy," *Journal of Pension Economics & Finance*, 2011, *10* (1), 75–97.

Brunello, Giorgio and Simona Comi, "The side effect of pension reforms on the training of older workers. Evidence from Italy," *The Journal of the Economics of Ageing*, 2015, *6*, 113–122.

Cahuc, Pierre, Francois Marque, and Etienne Wasmer, "A theory of wages and labor demand with intra-firm bargaining and matching frictions," *International Economic Review*, 2008, *49* (3), 943–972.

Card, David and Alan B Krueger, "Does school quality matter? Returns to education and the characteristics of public schools in the United States," *Journal of Political Economy*, 1992, *100* (1), 1–40.

_ , Francesco Devicienti, and Agata Maida, "Rent-sharing, holdup, and wages: Evidence from matched panel data," *Review of Economic Studies*, 2013, *81* (1), 84–111.

Carone, Giuseppe, Per Eckefeldt, Luigi Giamboni, Veli Laine, and Stéphanie Pamies Sumner, "Pension reforms in the EU since the early 2000s:

achievements and challenges ahead," *European Commission Discussion Paper 42*, 2016.

Carroll, Christopher D and Miles S Kimball, *Precautionary saving and precautionary wealth*, Springer, 2008.

Carta, Francesca and Marta De Philippis, "Working horizon and labour supply: the effects of raising minimum retirement age on middle-aged individuals," *Working paper*, 2019.

_ , Till M von Wachter, and Francesco D'Amuri, "Workforce aging, pension reforms, and firm dynamics," *Working paper*, 2019.

Cattaneo, Maria A, Chantal Oggenfuss, and Stefan C Wolter, "The more, the better? The impact of instructional time on student performance," *Education Economics*, 2017, pp. 1–13.

Cerdan-Infantes, Pedro and Christel Vermeersch, "More time is better: An evaluation of the full time school program in Uruguay," *World Bank Policy Research Working Paper 4167*, 2007.

Congressional Budget Office, "Options for reducing the deficit: 2017 to 2026," Technical Report, Congressional Budget Office 2016.

Contreras, Dante and Paulina Sepúlveda, "Effect of lengthening the school day on mother's labor supply," *The World Bank Economic Review*, 2016, *31* (3), 747–766.

Couch, Kenneth A and Dana W Placzek, "Earnings losses of displaced workers revisited," *American Economic Review*, 2010, *100* (1), 572–89.

COVIP, "La previdenza complementare. Principali dati statistici.," Technical Report, COVIP 2015.

Cribb, Jonathan, Carl Emmerson, and Gemma Tetlow, "Signals matter? Large retirement responses to limited financial incentives," *Labour Economics*, 2016, *42*, 203–212.

Davis, Steven J and Till M von Wachter, "Recessions and the costs of job loss," *Brookings papers on economic activity*, 2011, *2011* (2), 1–72.

de Grip, Andries, Maarten Lindeboom, and Raymond Montizaan, "Shattered dreams: the effects of changing the pension system late in the game," *The Economic Journal*, 2011, *122* (559), 1–25.

Dobbie, Will and Roland G Fryer Jr, "Getting beneath the veil of effective schools: evidence from New York City," *American Economic Journal: Applied Economics*, 2013, *5* (4), 28–60.

_ **and** _ , "The medium-term impacts of high-achieving charter schools," *Journal of Political Economy*, 2015, *123* (5), 985–1037.

_ **, Roland G Fryer, and G Fryer Jr**, "Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone," *American Economic Journal: Applied Economics*, 2011, *3* (3), 158–187.

Dominguez, Patricio and Krista Ruffini, "Long-term gains from longer school days," *IRLE Working Paper*, 2018.

Duggan, Mark, Perry Singleton, and Jae Song, "Aching to retire? The rise in the full retirement age and its impact on the social security disability rolls," *Journal of public economics*, 2007, *91* (7-8), 1327–1350.

Eide, Eric and Mark H Showalter, "The effect of school quality on student performance: a quantile regression approach," *Economics letters*, 1998, *58* (3), 345–350.

Engels, Barbara, Johannes Geyer, and Peter Haan, "Pension incentives and early retirement," *Labour Economics*, 2017, *47*, 216–231.

Eyles, Andrew and Stephen Machin, "The introduction of academy schools to England's education," *Journal of the European Economic Association*, 07 2018.

_ **,** _ **, and Sandra McNally**, "Unexpected school reform: academisation of primary schools in England," *Journal of Public Economics*, 2017.

Farber, Henry S, "Employment, hours, and earnings consequences of job loss: US evidence from the displaced workers survey," *Journal of Labor Economics*, 2017, *35* (S1), S235–S272.

Fetter, Daniel K and Lee M Lockwood, "Government Old-Age Support and Labor Supply: Evidence from the Old Age Assistance Program," *American Economic Review*, 2018, *108* (8), 2174–2211.

Figlio, David, Kristian L. Holden, and Umut Ozek, "Do students benefit from longer school days? Regression discontinuity evidence from Florida's sdditional hour of literacy instruction," *Economics of Education Review*, 2018, *67*, 171–183.

French, Eric, "The effects of health, wealth, and wages on labour supply and retirement behaviour," *The Review of Economic Studies*, 2005, *72* (2), 395–427.

_ **and John Jones**, "Public pensions and labor supply over the life cycle," *International Tax and Public Finance*, 2012, *19* (2), 268–287.

Giavazzi, Francesco and Michael McMahon, "Policy uncertainty and household savings," *Review of Economics and Statistics*, 2012, *94* (2), 517–531.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik instruments: what, when, why and how," *Working paper*, 2019.

Goodman, Joshua, "Flaking out: student absences and snow days as disruptions of instructional time," *NBER Working Paper Series*, 2014.

Grogger, Jeff, "Does school quality explain the recent black/white wage trend?," *Journal of Labor Economics*, 1996, pp. 231–253.

Gruber, Jonathan and David A Wise, *Social security programs and retirement around the world: the relationship to youth employment*, University of Chicago Press, 2010.

Hairault, Jean-Olivier, Thepthida Sopraseuth, and François Langot, "Distance to retirement and older workers 'employment: The case for delaying the retirement age," *Journal of the European Economic association*, 2010, *8* (5), 1034–1076.

Hamann, A Javier, "The reform of the pension system in Italy," Technical Report, International Monetary Fund 1997.

Hanel, Barbara and Regina T Riphahn, "The timing of retirement—New evidence from Swiss female workers," *Labour economics*, 2012, *19* (5), 718–728.

Hansen, Benjamin, "School year length and student performance: quasi-experimental evidence," *Available at SSRN 2269846*, 2011.

Hendren, Nathaniel, "Efficient welfare weights," *Working paper*, 2019.

Hincapie, Diana, "Do longer school days improve student achievement? Evidence from Colombia," *IDB Working Paper n. 679*, 2016.

Huebener, Mathias, Susanne Kuger, and Jan Marcus, "Increased instruction hours and the widening gap in student performance," *Labour Economics*, 2017.

Iacus, Stefano M, Gary King, and Giuseppe Porro, "Causal inference without balance checking: Coarsened exact matching," *Political analysis*, 2012, *20* (1), 1–24.

Inderbitzin, Lukas, Stefan Staubli, and Josef Zweimüller, "Extended unemployment benefits and early retirement: Program complementarity and program substitution," *American Economic Journal: Economic Policy*, 2016, *8* (1), 253–88.

INPS, "INPS- XV Rapporto Annuale," Technical Report, INPS 2016.

Jacobson, Louis S, Robert J LaLonde, and Daniel G Sullivan, "Earnings losses of displaced workers," *The American economic review*, 1993, pp. 685–709.

Jäger, Simon and Jörn Heining, "How substitutable are workers? Evidence from worker deaths," *Working paper*, 2019.

Jaravel, Xavier, Neviana Petkova, and Alex Bell, "Team-specific capital and innovation," *American Economic Review*, 2018, *108* (4-5), 1034–73.

Karlström, Anders, Mårten Palme, and Ingemar Svensson, "The employment effect of stricter rules for eligibility for DI: Evidence from a natural experiment in Sweden," *Journal of public economics*, 2008, *92* (10-11), 2071–2082.

Kline, Patrick and Christopher R Walters, "Evaluating public programs with close substitutes: The case of Head Start," *The Quarterly Journal of Economics*, 2016, *131* (4), 1795–1848.

Krueger, Alan B and Jörn-Steffen Pischke, "The effect of social security on labor supply: A cohort analysis of the notch generation," *Journal of labor economics*, 1992, *10* (4), 412–437.

Lalive, Rafael, Arvind Magesan, and Stefan Staubli, "Raising the full retirement age: default vs incentives," *Working Paper*, 2017.

Lammers, Marloes, Hans Bloemen, and Stefan Hochguertel, "Job search requirements for older unemployed: Transitions to employment, early retirement and disability benefits," *European Economic Review*, 2013, *58*, 31–57.

Lavy, Victor, "Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries," *The Economic Journal*, 2015, *125* (588), F397–F424.

_ , "Expanding school resources and increasing time on task: effects of a policy experiment in Israel on student academic achievement and behavior," *Journal of the European Economic Association*, forthcoming.

Lazear, Edward P, "Why is there mandatory retirement?," *Journal of political economy*, 1979, *87* (6), 1261–1284.

Lee, Jong-Wha and Robert J Barro, "Schooling quality in a cross–section of countries," *Economica*, 2001, *68* (272), 465–488.

Liebman, Jeffrey B, Erzo FP Luttmer, and David G Seif, "Labor supply responses to marginal Social Security benefits: Evidence from discontinuities," *Journal of Public Economics*, 2009, *93* (11-12), 1208–1223.

Liu, Regina Y et al., "Bootstrap procedures under some non-iid models," *The Annals of Statistics*, 1988, *16* (4), 1696–1708.

Mammen, Enno et al., "Bootstrap and wild bootstrap for high dimensional linear models," *The annals of statistics*, 1993, *21* (1), 255–285.

Manoli, Dayanand S. and Andrea. Weber, "The effects of early retirement age on retirement decisions," *Working Paper*, 2018.

Marcotte, Dave E, "Schooling and test scores: a mother-natural experiment," *Economics of Education Review*, 2007, *26* (5), 629–640.

_ **and Steven W Hemelt**, "Unscheduled school closings and student performance," *Education Finance and Policy*, 2008, *3* (3), 316–338.

Martins, Pedro S, Álvaro A Novo, and Pedro Portugal, "Increasing the legal retirement age: the impact on wages, workers flows and firm performance," *IZA Discussion Paper No. 1487*, 2009.

Mastrobuoni, Giovanni, "Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities," *Journal of public Economics*, 2009, *93* (11-12), 1224–1233.

Meroni, Elena Claudia and Giovanni Abbiati, "How do students react to longer instruction time? Evidence from Italy," *Education Economics*, Nov 2016, *24* (6), 592–611.

Mohnen, Paul, "The impact of the retirement slowdown on the U.S. youth labor market," *Working Paper*, 2019.

Mönckeberg B., Fernando, "Prevention of undernutrition in Chile experience lived by an actor and spectator," *Revista Chilena de Nutrición*, 2003.

Montizaan, Raymond, Frank Cörvers, and Andries De Grip, "The effects of pension rights and retirement age on training participation: Evidence from a natural experiment," *Labour Economics*, 2010, *17* (1), 240–247.

Munnell, Alicia H and Anqi Chen, "Trends in Social Security claiming," *Center for Retirement Research at Boston College Working Paper 15-8*, 2015.

Nguyen, Bang Dang and Kasper Meisner Nielsen, "The value of independent directors: Evidence from sudden deaths," *Journal of financial economics*, 2010, *98* (3), 550–567.

OECD, *Indicators of Employment Protection* 2015.

_ , "Dare to share: Germany's experience promoting equal partnership in families," 2016.

_ , "How is learning time organised in primary and secondary education?," 2016.

OECD, "Pensions at a Glance 2017: OECD and G20 indicators," Technical Report 2017.

Parinduri, Rasyad A, "Do children spend too much time in schools? Evidence from a longer school year in Indonesia," *Economics of Education Review*, 2014, *41*, 89–104.

Patall, Erika A, Harris Cooper, and Ashley Batts Allen, "Extending the school day or school year: A systematic review of research (1985–2009)," *Review of Educational Research*, 2010, *80* (3), 401–436.

Pérez, MA Ruz and Ángela Madrid Valenzuela, "Evaluación Jornada Escolar Completa (Resumen Ejecutivo)," *Santiago de Chile: Pontificia Universidad Católica de Chile, Dirección de Estudios Sociológicos*, 2005.

Pischke, Jörn-Steffen, "The impact of length of the school year on student performance and earnings: evidence from the German short school years," *The Economic Journal*, 2007, *117* (523), 1216–1242.

Pissarides, Christopher A, *Equilibrium unemployment theory*, MIT press, 2000.

Quigley, Timothy J, Craig Crossland, and Robert J Campbell, "Shareholder perceptions of the changing impact of CEOs: Market reactions to unexpected CEO deaths, 1950–2009," *Strategic Management Journal*, 2017, *38* (4), 939–949.

Rivkin, Steven G and Jeffrey C Schiman, "Instruction time, classroom quality, and academic achievement," *The Economic Journal*, 2015, *125* (588), F425–F448.

Rizzuto, Ronald and Paul Wachtel, "Further evidence on the returns to school quality," *Journal of Human Resources*, 1980, pp. 240–254.

Santiago, Paulo, Francisco Benavides, Charlotte Danielson, Laura Goe, and Deborah Nusche, *Teacher Evaluation in Chile 2013*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris, 2013.

Schmieder, Johannes F, Till M von Wachter, and Joerg Heining, "The cost of job displacement over the business cycle and its sources: evidence from Germany," *Working Paper*, 2018.

Seibold, Arthur, "Reference points for retirement behaviour: evidence from German pension discontinuities," *Working Paper*, 2019.

Sims, David P., "Strategic responses to school accountability measures: It's all in the timing," *Economics of Education Review*, 2008, *27* (1), 58–68.

Staubli, Stefan, "The impact of stricter criteria for disability insurance on labor force participation," *Journal of Public Economics*, 2011, *95* (9-10), 1223–1235.

_ **and Josef Zweimüller**, "Does raising the early retirement age increase employment of older workers?," *Journal of public economics*, 2013, *108*, 17–32.

Vestad, Ola Lotherington, "Labour supply effects of early retirement provision," *Labour Economics*, 2013, *25*, 98–109.

Vigtel, Trond Christian, "The retirement age and the hiring of senior workers," *Labour Economics*, 2018, *51*, 247–270.

Werning, Ivan, "Pareto efficient income taxation," *Working Paper*, 2007.

Woessmann, Ludger, "Schooling Resources, Educational Institutions and Student Performance: The International Evidence," *Oxford bulletin of economics and statistics*, 2003, *65* (2), 117–170.

_ , "The importance of school systems: evidence from international differences in student achievement," *Journal of Economic Perspectives*, September 2016, *30* (3), 3–32.

Wu, Chien-Fu Jeff et al., "Jackknife, bootstrap and other resampling methods in regression analysis," *the Annals of Statistics*, 1986, *14* (4), 1261–1295.