DISCUSSION PAPER SERIES

IZA

IZA DP No. 4187

Increasing the Legal Retirement Age: The Impact on Wages, Worker Flows and Firm Performance

Pedro S. Martins
Álvaro A. Novo
Pedro Portugal

May 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Increasing the Legal Retirement Age: The Impact on Wages, Worker Flows and Firm Performance

Pedro S. Martins
Queen Mary, University of London,
CEG-IST and IZA

Álvaro A. Novo
Banco de Portugal
and ISEGI, Universidade Nova de Lisboa

Pedro Portugal
Banco de Portugal,
Universidade Nova de Lisboa and IZA

Discussion Paper No. 4187
May 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 4187
May 2009

ABSTRACT

Increasing the Legal Retirement Age: The Impact on Wages, Worker Flows and Firm Performance*

Many pay-as-you-go pension systems have increased or plan to increase their legal retirement age (LRA) to address the financial consequences of ageing. Although the success of these policies is ultimately determined at the labour market, little is known about the effects of higher LRAs at the firm level. Here, we identify this effect by considering a legislative reform introduced in Portugal in 1994: women's LRA was gradually increased from 62 to 65 years while men's LRA stayed unchanged at 65. Using detailed matched employer-employee panel data and difference-in-differences matching methods, we analyse the effects of the reform in terms of a number of worker- and firm-level outcomes. After providing evidence of compliance with the law, we find that the wages and hours worked of older women (those required to work longer) were virtually unchanged. However, firms employing older female workers significantly reduced their hirings, especially of younger female workers. Those firms also lowered their output although not their output per worker.

JEL Classification: J14, J26, J63

Keywords: social security reform, older workers, matching estimators

Corresponding author:

Pedro S. Martins
School of Business and Management
Queen Mary, University of London
Mile End Road
London E1 4NS
United Kingdom
E-mail: p.martins@qmul.ac.uk

* We thank comments from John Addison, Hans Bloemen, Jonathan Gardner, Edwin Leuven, Pierre-Jean Messe and seminar participants at CAED (Chicago), *Universidade Nova de Lisboa*, IPEA (Brasilia), *Banco de Portugal*, *Université de Paris I*, *Universidade do Porto*, CEMFI (Madrid), Netspar (Utrecht), and INED (Paris). We also thank Lucena Vieira for computational assistance. Martins gratefully acknowledges research support from Netspar (The Netherlands). Opinions expressed in the paper do not necessarily reflect the views of *Banco de Portugal*. All remaining errors are of our responsibility.

1 Introduction

Many pay-as-you-go pension systems across the world have been under financial pressure due to the combined effects of increased life expectancy and lower fertility rates. Several countries have responded or plan to respond by adjusting the legal retirement age (henceforth, LRA) - the age at which workers are entitled to retire.[1] Moreover, the age of retirement is also likely to be suject to further changes due to legislation against age discrimination (e.g. the recent directives issued by the European Union) which may eventually lead to the abolishment of mandatory retirement ages.

Although changes in pension systems have typically been studied from the point of view of their labour supply consequences, there are several reasons why adjustments in retirement age can also affect firm behaviour and labour demand. For instance, firms may offer incentive schemes in which wages are below productivity when workers start their careers, and then gradually increase at a faster pace than productivity (Lazear 1979). In the context of these incentive pay structures, an *ex post* increase in the mandatory retirement age would be detrimental to firms' profitability, particularly in a context of downward wage rigidity or strict employment protection legislation.

Moreover, firms that are forced to retain workers for a period longer than initially expected may respond by decreasing their hirings of new staff as the older workers will only need to be replaced later. In this case, the net effect of higher mandatory retirement ages upon the sustainability of pay-as-you-go pension systems is weakened. On the one hand, social security payments made by workers forced to postpone their retirement will increase while pension outlays will fall. On the other hand, social security payments by workers who are not hired will presumably fall while unemployment benefits may increase.

While it has been established that economic incentives play an important role in retirement decisions (e.g. Meghir & Whitehouse (1997) find that increased earnings in work delay job exit while increased social security benefits delay the return to work), very little is known about the implications of changes in the mandatory age of retirement, particularly at the firm level. The only related paper we know is Ichino et al. (2007), which argues that increasing the retirement age helps solve pension problems only if the employment prospects of the elderly do not worsen.

[1] For instance, Schwarz & Demirgue-Kunt (1999) describe the cases of 17 countries that have increased their mandatory retirement ages between 1992 and 1998. See also Burtless & Quinn (2002) for an analysis of the U.S. case.

Their evidence suggests that, although displaced elderly workers initially lose out in terms of employment chances, later on there are no significant employability differences between older and younger displaced workers. However, an approach based on displacement cannot shed much light on the impact of increases in retirement age, when employment protection legislations are strict as is the case in many countries.[2]

Our evidence of the effects of higher retirement age is based on a quasi-experiment involving an increase in the legal retirement age: a law introduced in Portugal in 1993 which increased the mandatory retirement age of women from 62 to 65 years while leaving the corresponding age for men unchanged at 65. Moreover, instead of focusing only on the specific case of older workers, we pursue a broader analysis of the labour market, namely by considering several aspects of the personnel policies of firms. Finally, we also examine the consequences of the reform in terms of firm performance.

Specifically, we follow workers and their firms over time and compare them with suitable 'control' groups, based on detailed information about almost all wage earners and firms in the country. Using treatment effects methods, most notably a combination of the difference-in-differences and matching approaches (Heckman et al. 1997), we analyze the extent to which the extension of the legal retirement age changed the employment status, hours worked, and wages of women affected by legislative change. At the firm level, using similar methods, we study the effect of postponing the legal retirement age upon total hirings, separations, net job creation, and the hirings of different demographic groups. We also consider the effects upon firm performance.

In our main results, we find that the wages and hours worked of older women (those required to work longer) were virtually unchanged. However, firms employing old female workers significantly reduced their hirings, especially of young female workers. Those firms also lowered their output, although not their output per worker.

The remaining of the paper is organized as follows. Section 2 sketches the Portuguese pension system, before and after the new law. The econometric methodologies, including the construction of treatment and control groups, are described in Section 3. We then present the data in Section 4, while Section 5.1 measures the compliance with the new law. Finally, Section 5 presents the results and Section 6 concludes.

[2]See also Ashenfelter & Card (2002), who find that, in the US defined-contributions setting, the elimination of mandatory retirement for college and university faculty led to lower retirement rates.

2 The retirement law reform

As in many other countries, the pension system in Portugal is of the defined-benefit type, in which the amount of the pension awarded to an older individual depends on the number of years the individual worked and on some weighted average of the wages earned throughout the person's career. The amount of the pension therefore does not depend on the returns to financial assets over the period in which the worker made his or her contributions. Moreover, the funding of these pensions is typically carried out on a pay-as-you-go basis, in which current workers' contributions are used to pay the benefit of current retirees.[3]

In the early 1990s, the Portuguese pension system was facing the financial problems typical of defined-benefit pay-as-you-go systems. These problems arose due to population ageing, as a consequence of higher life expectancy and particularly low fertility rates. In 1993, for instance, those aged 65 and above corresponded to 21.6% of the working population (Banco de Portugal 1994). As a response to these circumstances, the Portuguese government decided in 1993 to raise the mandatory retirement age of women from 62 to 65 years ('Decreto-Lei 329/93'), thus equalising the LRA for men and women.[4]

The law indicated that the new age of retirement for women would be implemented gradually, presumably to smooth the impact upon the first cohorts of older women that would be affected when the reform came into force. Specifically, the retirement age for women increased by six months every year, until it converged in 1999 to the level of men (see Table 1). For instance, while a woman born on 31st December 1931 would be entitled to retire on 31st December 1993 (on her 62nd birthday), a woman born one day later, on 1st January 1932, would only be entitled to receive her pension on 1st July 1994 (when 62 years and six months old). However, due to the gradual phasing in of the new retirement age, women born six months later, on 1st July 1932, would reach retirement age on 1st July 1995, i.e. when 63 years old.

There are two additional aspects in the pension system in Portugal that need to be taken

[3]The main alternative type of pension systems are of the fully-funded, defined-contribution type, when benefits are based on the value of individual accounts to which workers contribute over their active lives. Defined-contribution pension systems tend to be riskier for individuals, as the value of the account will vary with fluctuations in interest rates. A new, hybrid system is the notional defined contribution type (Barr & Diamond 2006).

[4]The financial and insurance sectors were exempted from this provision of the law and are therefore removed from our empirical analysis. Moreover, the law also included other provisions, namely by making the formula that calculated the pension level less generous.

into account. The first is that a LRA denotes the age at which a worker is entitled to claim old-age pension, provided that the worker contributed to the social security system for a sufficiently long period. At that point in time, the labour contract established between the employer and the employee is automatically terminated. However, workers are free to sign new labour contracts, with the same or different employers. Moreover, earnings received from the new labour contract are not subject to any special taxation related to the amount received from the old-age pension.

The second aspect to be taken into account is that, as in many other countries, the social security legislation in Portugal allowed for some exemptions from the standard retirement age. Such exemptions, leading to early retirement, were typically observed for unemployed workers, workers in firms undergoing economic turbulence, and in jobs supposed to be particularly exhausting (e.g. air traffic controllers). These exemptions motivate our analysis of compliance (Section 5.1).

3 Identification and estimation

The feasibility of our evaluation exercise depends crucially on the suitability of the counterfactual groups that can be generated from the available data. We address this matter by carefully selecting units for the control group(s) and by using a combination of two methodologies typically proposed to tackle non-experimental settings: difference-in-differences and matching (Rosenbaum & Rubin 1983). In particular, we implement a difference-in-differences matching estimator (Heckman et al. 1997), which Smith & Todd (2005) show that may have the potential benefit of eliminating some sources of bias present in non-experimental settings, improving the quality of evaluation results significantly.

We take advantage of the characteristics of the dataset and of the new legal framework to construct treatment and control groups. In particular, we explore (i) the existence of data for the pre- and post-legislative periods; (ii) the source of variation that the gender-specific law introduced; and (iii) the availability of a rich set of covariates and of data originating from the same local labour market (Heckman et al. 1997).

In the limit, the new retirement law will have directly affected all women under the age of 62 and all firms that employed at least one such woman. Nonetheless, some specific groups of women were more likely to influence the firm's response to the new legal retirement age.

Such women include those who would have reached the legal retirement age in year $t + 1$ had the LRA remained at its value of year t. For instance, those aged [60; 60.5) by the end of 1992 would have presumably retired in 1994 under the previous age limit (62); however, due to the increase to 62.5 years, they will have had to postpone their retirement to 1995. We therefore assign such women (and/or their firms) to our treatment group.

The definition of the treatment group can also be extended beyond the first cohort of women affected by the new law. For example, women aged [55;60.5) by the end of 1992 had to postpone their retirement up to 1999, the year when the retirement age was equalized across genders. Finally, when analyzing the impact of the increase in the LRA on firm-level variables, we consider that a firm is a treated unit if it employs at least one of these women.

A related choice that needs to be made concerns the years that correspond to the 'before' and 'after' periods. Recall that, starting in 1994, the women's new LRA was increased each year by six months until it reached (the men's LRA of) 65 years in 1999. Two obvious candidates for our 'before' period are the years of 1992 and 1993. We choose 1992 because the new law was already under discussion in 1993, which may have prompted individuals and firms to react in anticipation. On the other hand, the government policy was unknow in 1992 and, therefore, that year should not suffer from any anticipation effects. Therefore, our main treatment group in the 'before' period includes all women aged [57.5; 60.5) (worker-level analysis) or firms with at least one woman in this age range (firm-level analysis) - see first column in Table 1.

Regarding the non-experimental control group, we adopt two definitions, depending on whether we are conducting an individual- or firm-level analysis. In the former case, we consider as our control the group formed by men in the same age group as the women included in the treatment group. As men's LRA was already 65 years when the new law came into effect over the 1994 to 1999 period, we can construct comparable control groups in this age-related dimension. Of course, this control group raises gender-related issues. These are, however, mitigated if we are willing to accept the time-invariance hypothesis of the D-in-D estimator (discussed in the next section). In other words, if the gender gap is constant over the analysis period, using men as control for women is less of an issue. Indeed, the data seems to support this hypothesis. Between 1991 and 1993, the log difference of worked hours between men and women was 0.098, 0.093 and 0.10, while the log difference of total remuneration was 0.39, 0.39

and 0.38; these values are also statistically (and economically) constant over time.

At the firm-level analysis (hirings, separations, net job creation and firm performance), we consider different control groups, based on whether firms employed in 1992 any women affected by the new law. For instance, one possible control group is made up of firms that do not employ any woman aged [60; 60.5) in 1992. In this case, the corresponding treatment group would be firms that employ at least one such woman in 1992. We then consider alternative treatment/control groups by broadening the range of ages that lead to the assignment of firms into each group. The broadest age range corresponds to firms that employ or do not employ any woman aged [55; 60.5) in 1992.

Besides any gender-related issues that may arise, the non-random assignment of the quasi-natural experiment may raise questions about selection into treatment status. These, as far as they are imputable to observables, can be handled by the matching methodology. To address differences between the two groups due to time-invariant non-observable factors, we combine both D-in-D and matching strategies (Heckman et al. 1997, 1998), the so-called D-in-D matching estimator. This and other aspects of our methodology are described in more detail in Appendix A.

4 Data

We use two datasets in our analysis. To study the issues of labour income, working hours, worker flows and firm performance, we use *Quadros de Pessoal*, a matched employer-employee panel data set. The impact on labour market transitions are analyzed with a quarterly employment survey, *Inquérito ao Emprego*.

4.1 *Quadros de Pessoal*

The main data source used in this paper is *Quadros de Pessoal* (QP), a longitudinal dataset matching firms and workers based in Portugal. The data are gathered every year by the Ministry of Employment and Social Security, based on a census of firms that employ at least one employee. Reported data cover all personnel working for the firm/establishment in a reference month (March, up to 1993, and October, from 1994).

Personnel on short-term leave (such as sickness, maternity, strike or holidays) are also included, whereas personnel on long-term leave (such as military service) are not reported.

Civil servants, the self-employed and domestic service are not covered, and the coverage of agriculture is low given its low share of wage-earners. Reported data include the firm's location, industry, employment, sales, ownership, legal setting, and the worker's gender, age, skill, occupation, schooling, hiring date, earnings, work duration, etc.

The mandatory nature of the survey leads to an extremely high response rate. Given the nature of the dataset, which covers not just every company with wage-earners, but also all of its workers, problems commonly faced by panel data sets, such as under- or over-sampling of certain groups and panel attrition, are much attenuated. Also, employer-reported wage information is known to be subject to less measurement error than worker-reported data.

Each firm entering the database is assigned a unique identifying number, in such a way that it can be followed over time. The Ministry of Employment implements several checks to ensure that a firm that has already reported to the database is not assigned a different identification number. Similarly, each worker also has a unique identifier, based on his/her social security number.

4.2 *Inquérito ao Emprego*

Our second dataset is taken from the nationally representative Portuguese quarterly employment survey, *Inquérito ao Emprego* (IE), conducted by *Instituto Nacional de Estatística*, the Portuguese statistics agency. We use data for the period 1992(2)-2000(4). In addition to employment status, the employment survey contains information on the individual's age, gender, schooling, etc.

The survey has a quasi-longitudinal nature: one sixth of the sample rotate out of the sample each quarter, so that we can track transitions from employment for up to five quarters. Transition rates are then obtained simply by identifying those employed individuals in the survey, who move out of employment over the subsequent quarter. In this paper, we distinguish between two destinations: unemployment and economic inactivity (i.e., withdrawal from the labour force).

The main restrictions imposed on the data set were that the individual be employed at the time of the survey, aged older than 55, and resident in mainland Portugal. Finally, due to potential sample attrition, we ensured that individuals appearing in subsequent surveys with the same identifier were in fact the same individual. The resulting sample size is 229,066 individuals. Among them we were able to identify 1,167 transitions from employment into inactivity.

5 Results

5.1 Measuring compliance

A common concern with the measurement of treatment effects is the effectiveness of the quasi-experiment. In other words, one needs to know how far reaching is the impact of the legislative change, particularly in a context in which early retirement is observed. In order to evaluate the effect of the new law on the labour force status of affected women we use the *Inquérito ao Emprego* data and specify conventional logit models to estimate the probability of being employed and the probability of being inactive. Based on time of the survey and on the age and gender of the individuals, we defined a dummy variable identifying the women likely to be affected by the change in the legislation. More specifically, this variable takes value one for women aged 62 to 62.5 years in 1994, for women aged 62 to 63 in 1995, and so on up to 1999, when the dummy is one for women aged 62 to 65. We called this variable the 'Treatment Group'.

The estimation results are provided in Table 2 where it can be seen that the probability of being employed for the treated group of women increased sizeably. According to the logit estimates, the odds ratio associated with the treatment group is 1.313, meaning that it is 31.3 percent more likely for a women affected by the increase of the retirement to be employed. Symmetrically, the probability of being inactive decreased significantly among the treated women, where the decline is estimated to be around 27.9 percent. The overall picture from these two logit regressions is that the new retirement age rules had a visible impact in the labour force status of affected women.

We also provide a more complete picture of the labour market changes that emerge from postponing the retirement age by looking at transitions out of employment. Since the Portuguese employment survey has a quasi-longitudinal nature, one can track transitions between labour market states for about five sixths of the sample. In particular, one can spot transitions from employment into inactivity among individuals who are old enough to consider retirement. Based on the age of the individuals, one should expect to see an increase in the hazard rate for women affected by the change in the legislation. This is indeed what is obtained from the estimation of a Cox proportional hazards model, where the time of the implementation of the new law (the 'After' variable in the specification) is treated as a time-varying covariate. The indication provided in Table 3 is that the hazard rate more than tripled among the affected women.

5.2 Worker-level analysis: effects on wages and working hours

In Table 4, we present a set of difference-in-differences matching (DDM) estimates for the effect of the treatment on the treated for total wages, working hours, and the probability of absence from work. The general result that emerges is that the impact of the increase of the mandatory age of retirement on these women's labour market outcomes are negligible. Neither income, nor working hours were affected by the extension of the working age. Also, the probability that a woman is an absentee, which could admittedly increase when requiring women to stay employed beyond their initial expectations, is not affected.

Before we discuss in more detail these DDM estimates, we shift our focus to the choice of the covariates used in the estimation of the propensity score and also to the plausibility of the assumption underlying the matching estimator. The choice of the variables in the specification of the probit model observed the basic principle that they should influence both the selection-into-treatment (to remain on the job) and the outcome variables. Thus, the variables included (see Table 5) are: potential experience and current job tenure and their quadratic terms, year dummies and (log) sales - to control for economy-wide and firm-specific shocks -, education level dummies, and sector of activity and regional dummies.[5] While the latter two sets of variables might influence more the outcome variable, clearly the other variables are simultaneously important in determining the decision to remain employed and the outcome variable.

The focus of Table 5 is, however, on the balancing properties of the matching procedure. For this purpose, we present a plethora of statistics, namely, the mean for the treatment and control groups for the unmatched and matched samples, the standardized bias measure suggested by Rosenbaum & Rubin (1985), and the joint significance tests and pseudo-R^2

[5] This table refers only to the propensity score matching procedure for the 'after' period. Similar testing schemes were conducted for the other components of the DDM estimator with overall results qualitatively identical. The full set of results is available from the authors upon request.

of the propensity score (probit model) estimation (Sianesi 2004). This table illustrates the importance of matching and its success. While before the kernel-based matching procedure the treatment and control groups exhibited clear differences (e.g. tenure differed by about 2 years), after matching these differences are reduced to statistical insignificance. This is also confirmed by the reduction obtained in the standardized bias and, finally, by the joint statistical significance of the covariates and by the pseudo-R^2 of the propensity score in the unmatched and matched samples estimation procedures. As it can be seen in the last two rows of Table 5, the pseudo-R^2 in the propensity score estimation that used only the treated units and the corresponding matched control units falls to values close to zero. The F-test complements this information, corroborating the view that matching has successfully eliminated any systematic observable differences between the treated and control groups.

With regards to the DDM estimates, we present two estimates, depending on the use of unbalanced panel data (which we treat as repeated cross sections) or balanced panel data. The researcher has typically these two options, and the choice of one over the other hinges on the question to be answered. In the present case, as women had access to early retirement schemes, one cannot exclude the possibility that, faced with unexpected extensions of their careers, some of them opted for such retirement schemes. Thus, by opting for the balanced panel data, we are in fact looking exclusively over time at those who (as expected by the legislator) extended their careers. For the present case, Table 4 reports these alternative estimates and both are statistically not different from zero.

To check on the sensitivity of our point estimates to the definition of the non-experimental control group, we consider two alternative definitions of control units. The obvious choice to compare women would be other women. This, however, raises difficulties in the current setting because all women younger than 62 years were affected by the new legislation. Therefore, we have one obvious choice - women older than 62 in 1993 - and a less obvious and, indeed potentially endogenous choice - younger women who did not have to postpone retirement in the 1994-1999 transition period. The last column of Table 4 presents the results. We find that the conclusions do not depend on the choice of the control group. Neither income, nor working hours were affected by the postponing of the retirement age. The same is true for the probability of being absent from work.

5.3 Firm-level analysis: effects on job and worker flows

We begin our firm-level analysis by studying the effects of the reform in terms of three main different labour market variables: hirings, separations, and net job creation. Net job creation in year t is defined as the difference in the total number of workers between year t and year $t-1$ in each firm. Hirings at year t are defined as the number of workers that are hired since year t up to year $t-1$. Separations are defined as the difference between hirings and net job creation.[6] Finally, in order to shed more light into any possible patterns resulting from the change in retirement age, we also decompose the total level of hirings into four groups of workers. These groups are defined according to the gender and the age (25 or younger; and older than 25) of the worker hired.

Because we are interested in understanding the net impact of the reform, all variables (hirings, separations and net job creation) are considered in a cumulative way when they refer to the 'after' period. Specifically, each variable results from summing the flow of the year under analysis and the flow of the same variable for all previous years in the treatment period up to that year under analysis. For instance, when we refer to the impact on hirings in 1997, we are comparing the sum of hirings in 1995, 96 and 97 in the treated group with the same sum for the control group.

Our models are estimated in within-firm differences, by taking as the dependent variable the difference between the value of the variable for each year and the value of the same variable in 1992. The matching method used is kernel matching.[7] We also impose the common support. The propensity score are estimated using a very large set of variables: a cubic in firm size (measured in terms of the number of workers), five dummies for firm size ranges, a quadratic in the share of women in the workforce, a cubic in the average total pay per worker, a cubic in the average total number of hours worked, a quadratic in the percentage of workers that are men aged 60 or more, the shares of voting rights held by domestic (private) and foreign investors, 57 industry dummies and 29 region dummies.

The sample is also restricted to firms with 100 or fewer employees, as large firms that employed older women were very difficult to match - it is difficult to find large firms that do

[6] Our method of counting worker flows based on annual data implies that we may underestimate hirings and separations, as we cannot track workers that are hired after the census month in year $t-1$ and that then separate before the census month in year t. However, these short-term flows are not important from our point of view in this paper.

[7] We have checked the robustness of the results using nearest neighbour matching and the results (available upon request) are very similar.

not employ at least one woman affected by the new LRA. This restriction eliminates only less than 2% of our sample, given the relatively low average size of firms in Portugal. We also checked that our results are robust to other cut-off thresholds (results available upon request).

We then consider three different periods over which we carry out this aggregation of flows: only 1995 (the first year that falls exclusively during the 'after' period), 1995 to 1997 and 1995 to 1999. These periods have been chosen in order to establish a correspondence between the different criteria that assign firms to either the treatment or the control groups. As mentioned before, strictly speaking all firms with at least one female employee in 1992 will be (directly) affected by the increase of the mandatory age of retirement, to the extent that, under the new law, such firms will be forced to retain those workers for a longer period than expected when the worker was hired.

Table 6 present the results concerning the impact of the higher LRA in terms of firm-level job and worker flows. In this and the following tables, the column indicating the period to which the estimate refers also indicates the criterion adopted to define the control and treatment groups. The correspondence between the period range examined and the definition of the treatment group is designed to allow us to study the impact on worker flows over the same period in which the law was binding in terms of preventing the older women employees from retiring. For instance, estimates for the period 1995-99 are also based on a treatment group made up of firms that employed in 1992 at least one woman aged 55 to 60 (while the control group corresponds to firms that employed zero women aged 55 to 60 in 1992). Similarly, estimates for 1995-97 are based on treatment group firms employing at least one woman aged 57-60 in 1992. Finally, estimates for 1995 only are based on treatment groups firms employing at least one woman aged 59-60 in 1992. Also, recall that each row corresponds to a separate estimate from a different matching analysis.

The main result that emerges from these tables is that hirings and separations fall significantly for treated firms. For instance, when considering the impact of the new law upon cumulative hirings over the 1995-99 period, the estimated effect is -1.26 (t-ratio of -3.02). A similar comparison in the case of cumulative separations, again over the 1995-99 period, indicates an estimate of -1.28. Given the similarity of the two effects, the cumulative net job creation effect is virtually zero.

Another important result from Table 6 is that wider periods of analysis and wider definitions of the treatment group translate into bigger effects. For instance, the effect on hirings goes from -0.52 in 1995 to -1.26 in 1995-99. In order to place these and other estimates in context, it is important to mention that, in each treated firm in 1992, the number of older women (old enough to assign the firm to the treatment group) is between around 1 (1995) and around 1.5 (1995-99 period). We conclude from this analysis that the new law had the effect of decreasing hirings by about one worker for each older worker retained in the firm.

In terms of separations, the range of relative effects (change in separations per retained worker) is broadly similar, again supporting the one-to-one relationship expected given the earlier evidence that the law was binding. Moreover, given the identity connecting hirings, separations and net job creation, the latter is hardly changed in the treatment group with respect to the control group, as documented in our results. As we mentioned above, we also find that, the longer the time range considered (e.g. 1995-1999 vs. 1995 only), the stronger the relative impact of the law in terms of decreased hirings, separations, or net job creation.

We also carry out balancing tests for our estimates, in order to check some of the assumptions underlying the matching method. We find that the matched sample leads to a much greater equality of the observables across the treatment and the control groups. In the few cases that the t-test of the equality of the means of the two groups is rejected at the standard levels of significance, the economic difference between the two groups is particularly small (results available upon request).

5.4 Firm-level analysis: effects on firm performance

As argued above, it is possible that firms' performance is negatively affected by the additional constraint imposed on their personnel policies when the LRA is increased. We assess this hypothesis by extending the framework we used for job and worker flows, considering now the impact of the treatment in terms of different measures of firm performance, namely sales and sales per worker.[8] Unlike before, each one of these variables is now measured in a single period (the last year of the range of years considered for the definition of the treatment/control groups).

We find - Table 7 - that there are relatively large and reasonably significant effects in terms of sales. The figures, across the different specifications and estimation methods, range

[8]We also consider sales net of the wage bill with very similar results, available upon request.

between -0.03 and -0.08 for the 1995 and 1995-1999 periods, respectively, always with t-ratios above 1.4. However, when considering the effects in terms of sales per worker, we find no significant differences between the treatment and the control groups (effects ranging from -0.02 for 1995 to 0 for 1995-1999).

5.5 Firm-level analysis: robustness

As mentioned above, we decompose the effect of the law upon hirings in terms of four different demographic groups (female workers aged 25 or less, male workers aged 25 or less, female workers aged 26 or more, male workers aged 26 or more) that may have been affected differently. The results are presented in Table 8. The results - Table 8 - indicate that the negative effect upon hirings is concentrated upon younger workers and, in particular, upon younger women. Bearing in mind that the total effect upon cumulative hirings from 1995 to 1997 was -1.26 (t-ratio of -3.02) - 6 -, the effect upon hirings of women aged 25 or less is of -0.45 (t-ratio of -3.23). On the other hand, the effect upon hirings of men aged 25 or less is only -0.29 (t-ratio of -2.83) and the effects upon women and men older than 25 are, respectively, -0.36 (t-ratio of -2.88) and -0.17 (t-ratio of -1.16).

With respect to the benchmark results on job and worker flows amd firm performance, the findings are also remarkably similar for different samples and matching methods. For instance, we also consider a different sample definition, of only firms present in all years since 1991 until 1999 (Tables 9 and 10). The advantage of this definition is that it rules out any possible impacts of compositional changes in the pool of firms analysed, as the same firms are followed over time, although at the cost the representativeness of the sample.

6 Conclusions

Increasing the mandatory retirement age has been considered an important policy to improve the financial sustainability of pay-as-you-go, defined-benefit pension systems in a context of population ageing. Although the success of any such policy is essentially determined at the labour market, our paper is the first to examine how firms adjust their personnel policies when forced to retain their older workers longer than initially expected.

We present quasi-experimental evidence on such response by firms, by examining the impact of a 1993 law in Portugal that increased the retirement age of women while leaving unchanged the retirement age of men. Using matched employer-employee panel data and difference-in-differences matching methods, we compare firms that, before the law was announced, employed women old enough to be either immediately or soon after affected by the new law with otherwise very similar firms but that did not employ any such women.

After checking that firms did indeed comply with the law, we find that the wages and the hours of the affected women were virtually unchanged. Moreover, we also find that 'treated' firms significantly reduced their worker flows (hirings and separations). In our preferred specifications, the results indicate that firms hire approximately one fewer worker for each older worker that is retained due to the higher mandatory retirement age. Moreover, we also find that younger workers and, in particular, younger women are the demographic groups most affected by the lower level of total hirings; and that firm sales fall but not sales per worker.

The result about fewer hirings suggests that the contribution of higher retirement ages to the sustainability of pensions may be weaker than previously assumed, at least over the short run.

References

Ashenfelter, O. & Card, D. (2002), 'Did the elimination of mandatory retirement affect faculty retirement?', *American Economic Review* **92**(4), 957–980.

Banco de Portugal (1994), Annual report, Technical report.

Barr, N. & Diamond, P. (2006), 'The economics of pensions', *Oxford Review of Economic Policy* **22**(1), 15–39.

Burtless, G. & Quinn, J. F. (2002), Is working longer the answer for an aging workforce?, Working Papers in Economics 550, Boston College.

Heckman, J., Ichimura, H., Smith, J. & Todd, P. (1998), 'Characterizing selection bias using experimental data', *Econometrica* **66**(5), 1017–1098.

Heckman, J., Ichimura, H. & Todd, P. (1997), 'Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme', *Review of Economic Studies* **64**(4), 605–654.

Ichino, A., Schwerdt, G., Winter-Ebmer, R. & Zweimuller, J. (2007), Too old to work, too young to retire?, IZA Discussion Papers 3110.

Lazear, E. (1979), 'Why is there mandatory retirement?', *Journal of Political Economy* **87**(6), 1262–1284.

Meghir, C. & Whitehouse, E. (1997), 'Labour market transitions and retirement of men in the UK', *Journal of Econometrics* **79**(2), 327–354.

Rosenbaum, P. & Rubin, D. (1983), 'The central role of the propensity score in observational studies for causal effects', *Biometrika* **70**(1), 41–55.

Rosenbaum, P. & Rubin, D. (1985), 'Constructing a control group using multivariate matched sampling methods that incorporate the propensity score', *American Statistician* **39**(1), 33–38.

Schwarz, A. & Demirgue-Kunt, A. (1999), 'Taking stock of pension reforms around the world', *World Bank, Social Protection Discussion Paper Series 9917* .

Sianesi, B. (2004), 'An evaluation of the Swedish system of active labor market programs in the 1990s', *Review of Economics and Statistics* **86**(1), 133–155.

Smith, J. & Todd, P. (2005), 'Does matching overcome LaLonde's critique of nonexperimental estimators?', *Journal of Econometrics* **125**(1-2), 305–353.

A Appendix - Econometric implementation

Let Y_{it}^{D} be the potential outcome of interest for individual i at time t had (s)he been in state D_i, where $D_i = 1$ if exposed to the program and 0 otherwise. Let treatment take place at time t. The fundamental identification problem lies in the fact that we do not observe, at time t, individual i in both states. Therefore, we cannot compute the individual treatment effect, $Y_{it}^{1} - Y_{it}^{0}$. One can, however, if provided with a convenient control group, estimate the average effect of the treatment on the treated.

The idea behind a D-in-D estimator is that we can use an untreated comparison group to identify temporal variation in the outcome that is not due to the treatment. However, in order to achieve identification of the general D-in-D estimator we need to assume

$$E[Y_{it}^{0} - Y_{it'}^{0} \mid D_i = 1] = E[Y_{it}^{0} - Y_{it'}^{0} \mid D_i = 0], \tag{1}$$

where t' is a time period before the program implementation. The assumption states that, over time, the outcome variable of treated individuals ($D = 1$), in the event that they had not been exposed to the treatment, would have evolved in the same fashion as actually observed for the individuals not exposed to the treatment ($D = 0$), known as the time invariance principle.

If the assumption expressed in (1) holds, the D-in-D estimate of the average treatment effect on the treated can be obtained by the sample analogs of

$$\widehat{\alpha}_{\text{D-in-D}} = \{E[Y_{it} \mid D_i = 1] - E[Y_{it} \mid D_i = 0]\} - \{E[Y_{it'} \mid D_i = 1] - E[Y_{it'} \mid D_i = 0]\}. \tag{2}$$

The time invariance assumption can be too stringent if the treated and control groups are not balanced in covariates that are believed to be associated with the outcome variable. The D-in-D setup can be extended to accommodate a set of covariates and this is usually done in a linear way, which takes into account eligibility specific effects and time/aggregate effects. In the following model, $\widehat{\alpha}_D$ corresponds to the D-in-D estimate obtained on a sample of treatment and control units

$$Y_{it} = \lambda D_i + \tau_t + \theta' Z_{it} + \alpha_D D_i \tau_t + \varepsilon_{it}, \tag{3}$$

where D_i is as before and represents the eligibility-specific intercept, defined over age and gender according to treatment rules, τ_t captures time/aggregate effects and equals 0 for the 'before' period and 1 for the 'after' period, and Z is a vector of covariates included to correct for differences in observed characteristics between individuals in treatment and control groups.

This estimator controls for both differences in the Zs and for time-specific effects, but it does not allow α_D to depend on Z and it does not impose common support on the distribution of the $Z's$ across the cells defined by the D-in-D approach (namely, before and after, and treatment and control). Additionally, this procedure might be inappropriate if the treatment has different effects for different groups in the population.

These pitfalls can be overcome by supplementing the D-in-D estimates with propensity score matching. The difference-in-differences matching (DDM) estimator adds to the simple D-in-D estimator the comparability on the observable covariates that characterizes the propensity score matching estimator.

The feasibility of the matching strategy relies on a rich set of observable individual characteristics, Z, to guarantee that the distribution of the individual characteristics important to each evaluation exercise is the same in the difference-in-differences cells. The matching process models the probability of participation and matches individuals with similar propensity scores. The time invariance assumption for the DDM estimator is

$$E[Y_{it}^0 - Y_{it'}^0 \mid p, D_i = 1] = E[Y_{it}^0 - Y_{it'}^0 \mid p, D_i = 0], \tag{4}$$

where $p = \Pr(D = 1|X)$ is the propensity score. When estimating the mean impact of the treatment on the treated the matching estimator requires a conditional mean independence assumption

$$E\left(Y_{it}^0|Z, D_i = 1\right) = E\left(Y_{it}^0|Z, D_i = 0\right) = E\left(Y_{it}^0|Z\right) \tag{5}$$

and also requires that there is a nonparticipant analogue for each participant which means that $\Pr(D = 1|Z) < 1$.

The DDM estimator takes two forms, depending on the nature of the data, namely, balanced panel data or repeated cross-sections. For the former case,

$$\widehat{\alpha}_{DDM} = E\left[(Y_t^1 - Y_{t'}^1) - \widehat{E}\left(Y_t^0 - Y_{t'}^0|P\right)\right], \tag{6}$$

where $\widehat{E}(Y|P)$ represents the expected outcome of individuals in the control group matched with those in the treatment group. In the case of the repeated cross-section, the DDM takes the form of

$$\widehat{\alpha}_{DDM} = E\left[Y_t^1 - \widehat{E}\left(Y_t^0|P\right)\right] - E\left[Y_{t'}^1 - \widehat{E}\left(Y_{t'}^0|P\right)\right], \tag{7}$$

where all variables are as above. We use both estimators.

Tables

Table 1: Treatment groups: Before and after the new retirement age

	Treatment groups by age sets (Before=1992)						
Year:	1992	1994	1995	1996	1997	1998	1999
LRA:	62	62.5	63	63.5	64	64.5	65
	[57.5, 58)						[64.5, 65)
	[58, 58.5)					[64, 64.5)	
	[58.5, 59)				[63.5, 64)		
	[59, 59.5)			[63, 63.5)			
	[59.5, 60)		[62.5, 63)				
	[60, 60.5)	[62, 62.5)					

Notes: (1) Treatment group: The set of individuals (women) who would have retired in year t if the legal retirement age (LRA) had remained at its value of year $t-1$. For example, women in the age group [60, 60.5) in 1992 would have retired in 1994 if the LRA had remained at 62 years; (2) Before period: the 'before' is always set to 1992, when the women's LRA was 62 years and no legislative change was expected.

Table 2: labour Force Status: 1992-2000 or 1992 and 2000 only (Logit results)

Regressor	Labour Force Status: Employment	Inactivity	Labour Force Status: Employment	Inactivity
Gender (Female=1)	-1.242	1.333	-1.289	1.355
	(0.010)	(0.010)	(0.021)	(0.021)
Age Group	-0.047	0.072	-0.108	0.139
	(0.028)	(0.028)	(0.061)	(0.061)
Treated Group	0.272	-0.327	0.389	-0.420
	(0.031)	(0.031)	(0.067)	(0.068)
Number of observations	229,066	229,066	49,701	49,701
Wald test	28,553.6	31,525.4	6,158.3	6,636.2

Source: *Inquérito ao Emprego*. The specification includes 17 age and 8 year dummies (pr one year dummy in the case of the last two columns). Standard errors in parenthesis.

Table 3: Transition from Employment into Inactivity (Cox Hazard Model with Time-Varying Covariates)

Regressor	
Female	-0.309
	(0.069)
Age Group	0.338
	(0.298)
Age Group × Female	-0.311
	(0.177)
Age Group × After	-0.220
	(0.210)
Age Group × After × Female	0.724
	(0.309)
Number of observations	1,167
Wald test	47.3

Source: *Inquérito ao Emprego*. The specification includes 8 year dummies. Standard errors in parenthesis.

Table 4: Labour market outcomes: Impact on postponed women retirees' total income, working hours and probability of working

Control group	Men				Women 62+		Women 50-55	
			Matching		Matching		Matching	
Variable	unrest.$^{(1)}$	rest.$^{(2)}$	c. sect$^{(3)}$	panel$^{(4)}$	c. sect.	panel	c. sect.	panel
Log earnings	0.008	-0.015	0.005	0.008	0.029	-0.023	-0.005	-0.003
	(0.013)	(0.010)	(0.015)	(0.011)	(0.015)	(0.022)	(0.015)	(0.011)
	52,120	52,120	53,570	10,204	38,067	4,953	53,586	6,788
	0.08	0.44	-	-	-	-	-	-
Log hours	-0.033	-0.028	-0.026	0.009	0.019	0.006	-0.026	-0.010
	(0.007)	(0.007)	(0.010)	(0.009)	(0.013)	(0.016)	(0.010)	(0.009)
	50,628	50,628	52,166	9,823	39,274	5,041	52,178	6,850
	0.04	0.08	-	-	-	-	-	-
Pr(Absentee)$^{(5)}$	-	-	-0.011	-	-0.015	-	-0.011	-
	-	-	(0.009)	-	(0.009)	-	(0.009)	-
	-	-	66,811	-	57,114	-	66,815	-
	-	-	-	-	-	-	-	-

Notes: The values reported for each pair variable and estimator are point estimate, standard error, number of observations and R^2. (1) The D-in-D unrestricted estimator does not control for confounding factors; (2) The OLS D-in-D restricted estimator is based on a linear specification, controlling for observable characteristics; (3) DDM estimator with kernel matching on the propensity score with repeated cross-section data; (4) DDM estimator with kernel matching on the propensity score with balanced panel data. The set of variables used with the estimation of the propensity score and in the restricted OLS D-in-D estimator are reported in Table 5. (5) It refers to the probability that a employee although registered in QP is reported as having worked zero hours, and (s)he is taken as absentee.

Table 5: Balancing properties of the kernel based propensity score matching for the unbalanced panel data in the after period

Variable	Sample	Mean Treated	Control	After t-test p-value$^{(1)}$	% bias$^{(2)}$	Reduction |bias|
		Unbalanced panel data (as repeated cross-section)				
Experience	Unmatched	52.44	52.19	0.000		
	Matched	52.43	52.34	0.163	3	62.6
Experience^2	Unmatched	2759.20	2734.40	0.000		
	Matched	2758.40	2749.30	0.162	3	63.2
Tenure	Unmatched	15.58	17.80	0.000		
	Matched	15.60	15.51	0.704	0.8	95.6
Tenure^2	Unmatched	381.64	484.89	0.000		
	Matched	382.44	382.51	0.994	0	99.9
Total sales	Unmatched	7.02	7.78	0.000		
	Matched	7.03	7.09	0.227	-2.5	92.1
Education:						
High school	Unmatched	0.03	0.03	0.093		
	Matched	0.03	0.04	0.739	-0.8	73.9
College	Unmatched	0.03	0.04	0.001		
	Matched	0.03	0.03	0.402	-1.7	71.2
Year dummies:						
1994	Unmatched	0.17	0.18	0.191		
	Matched	0.17	0.16	0.471	1.5	33.6
1995	Unmatched	0.20	0.20	0.603		
	Matched	0.20	0.20	0.765	-0.6	29.5
1996	Unmatched	0.14	0.15	0.006		
	Matched	0.14	0.14	0.746	-0.7	86
1997	Unmatched	0.18	0.17	0.071		
	Matched	0.18	0.18	0.824	0.5	84.6
1998	Unmatched	0.15	0.14	0.246		
	Matched	0.15	0.15	0.818	-0.5	75.1
1999	Unmatched	0.17	0.16	0.107		
	Matched	0.17	0.17	0.909	-0.2	91.1

Observations:		
On common support	4,324	13,259
Off common support	12	0
	Unmatched	Matched
|Bias| summary statistics:		
Mean	8.82	1.01
Std. Dev.	11.34	0.95
Maximum	56.93	3.70
Minimum	0.34	0.01
Pseudo $R^{2(3)}$	0.135	0.002
Joint F-test, p-value	0.000	0.998

Notes: The table does not exhaustively list all variables included in the probit model used to estimate the propensity scores; we omit from the table the balancing property of sector of activity and regional dummy variables. (1) The p-value of the t-test for the equality of means in the treated and control groups, both before and after matching. (2) Bias is the standardized bias as suggested by Rosenbaum & Rubin (1985) reported together with the achieved percentage reduction in |bias|. (3) Pseudo R^2 from the probit model estimation of the propensity scores, including all variables reported above, before and after the matching process (Sianesi 2004).

Table 6: Effects on Flows, Probit Pscore, Pooled Data

	Year	*ATT*	*t(ATT)*	Size	*Treated*	*Control*
Hirings						
	1995	-0.52	-2.27	21.1	5870	59051
	1995-97	-0.67	-2.36	21.8	4937	48172
	1995-99	-1.26	-3.02	22.0	4409	43025
Separations						
	1995	-0.81	-3.57	21.1	5870	59051
	1995-97	-1.04	-3.27	21.8	4937	48172
	1995-99	-1.28	-2.85	22.0	4409	43025
Net job creation						
	1995	0.30	1.37	21.1	5870	59051
	1995-97	0.37	1.19	21.8	4937	48172
	1995-99	0.02	0.07	22.0	4409	43025

Notes: Source: Own calculations based on *Quadros de Pessoal*. *ATT* refers to the average treatment effect in terms of the worker flow considered and at the year under analysis. *t(ATT)* denotes analytical standard errors. Estimation is carried out using fixed effects. Estimates are based on the difference in the accumulated level of the worker flows from 1995 until the year under analysis and the base year, 1992. *Treatment* and *Control* indicates the number of firms assigned to the treatment and control groups respectively.

Table 7: Effects on Sales, Probit Pscore, Pooled Data

	Year	*ATT*	*t(ATT)*	Size	*Treated*	*Control*
Sales						
	1995	-0.03	-1.77	21.4	5165	50694
	1995-97	-0.03	-1.45	21.6	4365	42468
	1995-99	-0.08	-3.39	22.1	3974	38187
Sales per worker						
	1995	-0.02	-1.30	21.4	5165	50694
	1995-97	0.00	0.07	21.6	4365	42468
	1995-99	0.00	0.13	22.1	3974	38187

Notes: Source: Own calculations based on *Quadros de Pessoal*. *ATT* refers to the average treatment effect in terms of the worker flow considered and at the year under analysis. *t(ATT)* denotes analytical standard errors. Estimation is carried out using fixed effects. Estimates are based on the difference in the accumulated level of the worker flows from 1995 until the year under analysis and the base year, 1992. *Treatment* and *Control* indicates the number of firms assigned to the treatment and control groups respectively.

Table 8: Effects on Hires by Demographic Group, Probit Pscore, Pooled Data

	Year	*ATT*	*t(ATT)*	*Treated*	*Control*	
Younger men						
	1995	-0.12	-1.83	21.1	5870	59051
	1995-97	-0.09	-1.33	21.8	4937	48172
	1995-99	-0.29	-2.83	22.0	4409	43025
Younger women						
	1995	-0.24	-2.88	21.1	5870	59051
	1995-97	-0.26	-2.76	21.8	4937	48172
	1995-99	-0.45	-3.23	22.0	4409	43025
Older men						
	1995	0.00	0.00	21.1	5870	59051
	1995-97	-0.12	-1.48	21.8	4937	48172
	1995-99	-0.17	-1.16	22.0	4409	43025
Older women						
	1995	-0.16	-1.99	21.1	5870	59051
	1995-97	-0.21	-1.62	21.8	4937	48172
	1995-99	-0.36	-2.88	22.0	4409	43025

Notes: Source: Own calculations based on *Quadros de Pessoal*. *ATT* refers to the average treatment effect in terms of the worker flow considered and at the year under analysis. *t(ATT)* denotes analytical standard errors. Estimation is carried out using fixed effects. Estimates are based on the difference in the accumulated level of the worker flows from 1995 until the year under analysis and the base year, 1992. *Treatment* and *Control* indicates the number of firms assigned to the treatment and control groups respectively.

Table 9: Effects on Flows, Probit Pscore, Only Continuing Firms

	Year	*ATT*	*t(ATT)*	Size	*Treated*	*Control*
Hirings						
	1995	-0.70	-3.17	21.7	3442	34483
	1995-97	-0.94	-2.77	21.7	3442	34483
	1995-99	-0.71	-2.25	21.7	3442	34483
Separations						
	1995	-0.84	-4.12	21.7	3442	34483
	1995-97	-0.95	-4.18	21.7	3442	34483
	1995-99	-0.78	-3.02	21.7	3442	34483
Net job creation						
	1995	0.14	0.56	21.7	3442	34483
	1995-97	0.02	0.05	21.7	3442	34483
	1995-99	0.07	0.21	21.7	3442	34483

Notes: Source: Own calculations based on *Quadros de Pessoal*. *ATT* refers to the average treatment effect in terms of the worker flow considered and at the year under analysis. *t(ATT)* denotes analytical standard errors. Estimation is carried out using fixed effects. Estimates are based on the difference in the accumulated level of the worker flows from 1995 until the year under analysis and the base year, 1992. *Treatment* and *Control* indicates the number of firms assigned to the treatment and control groups respectively.

Table 10: Effects on Sales, Probit Pscore, Only Continuing Firms

	Year	*ATT*	*t(ATT)*	Size	*Treated*	*Control*
Sales						
	1995	-0.05	-2.12	21.7	2800	28026
	1995-97	-0.05	-2.02	21.7	2800	28026
	1995-99	-0.08	-3.33	21.7	2800	28026
Sales per worker						
	1995	-0.04	-1.96	21.7	2800	28026
	1995-97	-0.02	-1.04	21.7	2800	28026
	1995-99	-0.03	-1.15	21.7	2800	28026

Notes: Source: Own calculations based on *Quadros de Pessoal*. *ATT* refers to the average treatment effect in terms of the worker flow considered and at the year under analysis. *t(ATT)* denotes analytical standard errors. Estimation is carried out using fixed effects. Estimates are based on the difference in the accumulated level of the worker flows from 1995 until the year under analysis and the base year, 1992. *Treatment* and *Control* indicates the number of firms assigned to the treatment and control groups respectively.